A Study of Agent Influence in Nested Agent Interactions

Thesis

A study of agent influence in nested agent interactions

A dissertation submitted in partial satisfaction of the requirements
for the degree of Doctor of Philosophy in Computer Science

Robert Logie
Faculty of Mathematics, Computing and Technology
The Open University
Milton Keynes
England

June 2009

DATE OF SUBMISSION: 30 JUNE 2009
DATE OF AWARD: 18 JUNE 2010

Dedication

For Hiroko

Acknowledgements

I am greatly indebted to my supervisors, Jon G. Hall and Kevin G. Waugh, for their advice, support and encouragement throughout this work.

Thank you Janet van der Linden of the Open University for support outside of the research loop.

Thank you to my wife, Hiroko, for putting up with me throughout this work.

Thank you to my family and friends for their support and encouragement.

Thank you to Osaka Gakuin University and my colleagues in the faculty of informatics for their support and encouragement.

This work has involved a lot of travel and would like to thank Dr. Laura Davey and Dr Stefano Castelvecchi for arranging accommodation for me on many occasions.

I would like to thank the Japanese Ministry of Foreign Affairs for sending my wife to Europe for a short period: the move allowed me to start this work.

Related publications

Some the preliminary research leading to this work has been published.

Reactive food gathering

This is an investigation into emergent behaviour and the first steps at examining how deontic notions may be applied to emergent behaviour and simple reactive agents.

R. Logie, J. G. Hall and K. G. Waugh, Reactive food gathering. in F. Toni and P. Torroni, editors, *CLIMA VI, volume 3900 of Lecture Notes in Computer Science*, pages 406 – 413. Springer, 2005.

Towards mining for influence in a multi agent environment

One of the early threads in this research was using data mining to identify other agent influence. This paper details a characterisation of the problem from a data mining perspective.

R. Logie, J. G. Hall and K. G. Waugh, Towards mining for influence in a multi agent environment. in P. Abraham, editor, *IADIS European Conf. Data Mining*, pages 97 – 101. IADIS, 2008.

Abstract

This work develops a theory of agent influence and applies it to a coached system of simple reactive agents. Our notion of influence is intended to describe agent ability which is contingent on the actions of other agents and we view such behaviours as being "nested". An agent may have the ability to make A hold only if another agent has carried out a particular action. Our analysis of this is based on a combination of the observation of the effects of an agent's actions in a bounded environment and observations on what may be changed in that environment and is intended to allow for a logical representation of nested behaviours. We build on this notion to develop a theory of influence which we offer as an extension of existing systems for representing agency and its effects.

The notion of an agent being able to "see to it" that something is brought about has been a useful device for reasoning about agent ability. These so-called STIT semantics have been developed by a number of researchers. Standard STIT semantics allow statements of the form $[\alpha\ stit\colon A]$ which says that agent α has the ability to see to it that A holds. Although based on the concept of agent action STIT semantics also allow for the representation of concepts involving what may be thought of as *inaction*. An agent deciding, for example, not to execute a particular action may be characterised as seeing to it that it does not see to it that A, $[\alpha\ stit\colon [\alpha\ stit\colon \neg A]]$. STIT encourages nesting and although this nesting extends across actions *within* an agent it does not extend easily *across* agents. So called *other agent* statements of the form $[\beta\ stit\colon [\alpha\ stit\colon A]]$ do not make sense in standard STIT semantics because β seeing to it that α sees to it that A holds implies that β has some dominion over α which, in turn, compromises α*'s* agency. Although the statement makes no sense under standard STIT it does make sense in an intuitive way and Brian Chellas [31] notes that it would be:

"...bizarre to deny that an agent should be able to see to it that another agent sees to something"

This is also mentioned in Belnap et al. [8, page 275]. Chellas is correct and there are numerous settings in which other agent STIT does make sense. These settings, which are captured in various readings of STIT, may bring a great deal of system level overhead. In a normative system, for example, β may have the option of imposing a sanction on α if α fails to bring about A and in this sense may be thought of as seeing to it that α sees to it that A holds. Similarly a deontic reading may place β in a position where it is able to place an obligation on α to bring about A. These readings allow for sensible interpretation of other agent STIT but the examples above require that agents have sufficient awareness of personal utility be able to manage sanctions or that they are able to reason about obligations. These readings offer nothing for simple agents with limited resources and abilities.

We offer another reading for the STIT element, one based on the concept of agent influence and one which carries minimal system level overhead. Because influence may be contingent on simultaneous or sequential behaviour by a number of agents it is extendible across agents and offers a means of addressing other agent statements. We extend the standard STIT semantics of Horty, Belnap and others with the introduction of "leads to" and "may lead to" operators which allow us to move our analysis into a setting where observation provides evidence of influence. We then explore the manifestation of influence in a number of scenarios. After exploring how influence manifests itself we then offer a partial logical characterisation of the influence operators and discuss its relationship with standard STIT.

Building on these semantics and the partial logical characterisation we then explore the practical use of our theory of influence in an agent learning system. We describe experiments with a system specified by safety and liveness properties and having two broad classes of agents, actors and coaches. Actor agents will manipulate their environment and coaching agents will observe the actor's behaviour and its effects using aggregated observations to generate new behaviours which are then seeded in the environment to modify actor behaviour.

We then offer a discussion and evaluation of our theory and its applications indicating where it may be further developed and applied.

Contents

List of Hypotheses

List of Observations

List of Definitions

List of Tables

List of Figures

Nomenclature

Abbreviations and acronyms

ANTS	Autonomous Nano-Technology Swarm
$\mathcal{L}_{TD}$	Logic of time division
AGM	The Alchourrón, Gärdenfors and Makinson theory of belief revision
ATL	Alternating Time temporal logic
BDI	The Beliefs, Desires and Intentions agent model
BOID	An abstract agent architecture which extends BDI by considering agent obligations, Beliefs, Obligations, Intentions, Desires
COIN	Collective Intelligence
CPS	co-operative problem solving
DAI	Distributed Artificial Intelligence
DDM	Distributed Data Mining
ESSLLI	The European Summer School in Logic, Language and Information
KD	One of a family of standard modal logics

$KD45$	One of a family of standard modal logics
$MABS$	Multi Agent Based Simulation, an agent system formalism using a different notion of agent influence
$NMAS$	Normative Multi Agent System, an architecture based on speech act theory
NoA	an abstract agent architecture for Normative Agents
PCR	The ratio of positive to counter evidence for a hypothesis
PMC	The difference between positive and counter evidence tallies for a hypothesis
TMS	Truth Maintenance System

Logic and other operators

$\Box$	Standard modal necessity operator
$\Diamond$	Standard modal possibility operator
$\vdash$	Is a theorem
$\leadsto$	Influence leads to operator, developed here
$\Diamond\!\!\leadsto$	Influence may lead to operator, developed here
$\models$	Is a model of
$\rightarrow$	Logical implication
F	Temporal operator, in the future ...
P	Temporal operator, in the past ...

Modal logic axioms

4	What is necessarily so is necessarily necessarily so
4_{agent}	4 with agent extension
C	Convergence axiom
C_{agent}	Convergence axiom with agent extension
N	Rule of necessitation
RE	Rule of equivalence
RE_{agent}	Rule of equivalence with agent extension
S	Syllogism
T	Tautology

Modal logic rules

$\overline{N}$	Negative necessitation
RK	Rule of inference
RN	Rule of necessitation
RR	Rule of inference

Modal logic validities

$Df\Diamond$	What is possible is just what is not necessarily not

Chapter 1

Introduction – nested other agent ability

The notion of agency is not new, in computational terms, but much of the research on agency has been focused on agents rather than systems and societies of agents. This is a natural focus, if one sees an agent metaphorically as an individual. This metaphor is evident in many of the research areas branching from or supporting agent research. Normative systems focus on the effects of norms on the behaviour of individual agents. Agent communications languages focus on agent mental states [108]. These are both implicitly single agent concepts which may be aggregated and viewed at a system level but this does not necessarily mean that they may be aggregated and viewed as a *societal* concept.

We are interested in social rather than individual agency and to this end the main thread through this work is the investigation of a logical representation of nested behaviours. The nesting brings both social requirements and social benefits. The benefits are that a group of agents acting together or in sequence may have greater ability to exert influence on its environment than individual agents. The requirements are that in order to exert this greater influence agents must operate within a social context that allows the group – rather than individuals – to maximise influence on an environment. In order to investigate social agency we consider behaviour in agent agnostic terms, this makes sense as it allows us to examine behaviours and how agents contribute to them rather than agents and how they contribute to behaviours. Our view of agency is framed in so called STIT semantics, STIT is an acronym like contraction of *s*eeing *to it* and characterises agents in terms of their abilities to see to it that some proposition, the complement of a STIT expression, holds true. In this context agents are simply mechanisms for exercising ability. Their agency is a result of their having choices but their internal mechanisms and, indeed, their mechanisms for bringing about propositions, are of no concern. Standard STIT is expressive enough to deal with single agents and its semantics encourage nesting. For single agents this nesting provides a convenient way of expressing acting and refraining from acting – an agent may see to it that it does not see to it that A holds. Such nesting runs into difficulties when dealing with societal nesting, nesting where other agents are involved. If two agents, α and β, may act simultaneously to bring about A or they may act sequentially so that one agent opens the possibility for the other agent to bring about A. The former presents no great difficulties but the latter, the case of an *other agent nested* STIT, where β sees to it that α sees to it that A holds is much more complex. Intuitively this makes sense, when Chellas first proposed STIT operators he did not state postulates for multiple agents (see Xu [125]) but Chellas [31] (and mentioned in Belnap et al. [8, page 275]) notes, in a statement which we have already met in the abstract and shall call upon again, that it would be:

> "...bizarre to deny that an agent should be able to see to it that another agent sees to something"

The intuitive sensibilities of this nesting are not easily represented within the more formal constraints of STIT theory and they present considerable difficulties. Our contention is that by replacing the "strict" sees to it reading with an alternative "influences" reading, one which allows for the fact that an agent's abilities may be contingent, we may sensibly characterise such nested statements in a formal setting. We contend that "influences" is a valid reading of "sees to it" that captures dynamic aspects of interactions with an environment and other agents and that where an agent may genuinely and unambiguously see to it that A holds then "sees to it" and "influences" are equivalent. We contend, further, that in a practical setting coaching agents which attempt to maximise the influence of other actor agents will be able to constrain agent paths through potentially enormous state spaces in such a way that speeds the discovery of complex patterns of agent interaction which influence the agent environment.

1.1 Thesis outline

We start with a survey of the related literature in chapter 2. In chapter 3 we explore our notion of agent influence and develop this into a practical theory in chapter 4. We present a partial logical characterisation of our theory in chapter 5 so as to indicate its compatibility with standard STIT theory. Practical aspects are considered in chapters 6 and 7 where we explore how a coaching agent may detect influence and analyse it in such a way as to synthesise appropriate behaviours for simple reactive agents. Chapter 8 discusses agent influence and its performance in a practical setting and closes the work by discussing future avenues of research and possible applications.

The contribution of this thesis begins with an exploration of influence then moves on to develop a practical theory allowing us to formalise our observations. We then consider logical aspects of our theory of influence and offer a partial logical characterisation to support our claims that influence may be used as an alternative reading of strict STIT and that in certain conditions may be seen to be equivalent to strict STIT. We then return to agents and agent systems to consider how a coaching agent situated in the same environment as a number of actor agents may use its observations in conjunction with our theory of influence to synthesise new and more influential patterns of agent behaviour.

The areas explored are outlined briefly below.

1.1.1 An exploration of influence

We described our notion of influence as an *alternative* reading for STIT, there are already other alternatives and these others provide mechanisms for addressing the same difficulty. Belnap et al. [8, page 271] list deontic, disjunctive, probabilistic and strategic readings as alternatives and each of these goes some way to addressing the difficulties of nested STIT. They also bring additional difficulties, a deontic reading, for example, brings a requirement for a system of transferring obligations between agents and has an attendant requirement for a system of sanctioning agents that fail to honour obligations. Obligations and sanctions bring additional systems and additional complexity making their application to systems of simple agents difficult. The strategic reading is closest to our approach but is not as general as our notion of influence.

We have already mentioned influence several times, what exactly is influence? Chellas's statement, above, is a recurring theme in this work. Milner [87] provides us with another touchstone which recurs throughout this work, he notes that:

> The behaviour of a system is exactly what is observable.

Our view of influence is based on observed links between an agent's choice of action and any changes in that agent's environment. Described simplistically, if an agent's choice of an action has influence in a set of circumstances then we will observe consistent results in the environment following from this choice. If there are preconditions, perhaps another agent's action or the presence of another object, then we will also observe inconsistent results. If the agent has no influence in a dynamic environment then we will observe inconsistent results. Our initial exploration of influence is driven partially by experiment and is concerned with identifying single agent influence by observation and distinguishing genuine influence from environment noise. We then extend our single agent observations to groups of agents acting either in parallel or serially.

The fact that our notion of influence is based on observation gives it a flexibility that is absent from other readings. In a simple system, one where it is only the laws of physics which govern agent influence, then our observations will be of a process of cause and effect. If the system is complex and has societal laws which support obligations and sanctions then cause and effect may not be straightforward but our hypotheses of influence will cope with this by identifying behaviour sequences which lead to propositions being brought about. We are interested in raw evidence of influence and not the underlying mechanisms of influence.

Our investigation of influence leads us to a system of hypothesising about agency which frees us from detailed consideration of agents and examines their societal abilities to look for ways of producing agent communities that express greater influence than do their individuals. This leads us to the notion of societal gateways through which behaviour passes as influence expands. We develop agents and agent societies to a point where gateways are determinable, and show how to synthesise behaviours of communities that can move through the gateways. The notion of gateways rises from a novel view of the partitioning of a system's state space into domains of influence characterised by single agent or multiple agent involvement.

The societal view gives us insights into aspects of obligation and ability. This is most notable in the area of system specification and the "ought implies can" deontic identity. If a system is intended to carry out a task then it must be able to do so in some manner and this may not be immediately evident. Our approach may admit a characterisation of this identity by way of sequences of gateways in a system's state space and in doing so extends the semantic reach of standard STIT theory.

1.1.2 A practical theory of influence

We have already stated that this is an agent agnostic approach to agency. Complete agnosticism – a black box agent – is not really viable, we have already indicated that agent influence is driven by an agent's choice of action. Here we outline an agent algebra based on observable behaviours. Observability here may require some cooperation from agents, an agent's choice mechanism is of no interest but its choices are. It is not unreasonable to expect these choices to be observable even after the fact, an agent may simply report "I did this..".

We base our algebra on agent choice and develop a notation which allows us to carry our hypotheses of influence into a modal logic, branching time framework. This gives us a representation of agents as condensed binary choice mechanisms which map onto our theory of influence and put us on a position to consider the logical underpinnings of agent influence.

1.1.3 A partial logical characterisation of influence

We introduce two modal operators, "leads to" and "may lead to", which fit with our notion of influence and provide an intermediary between the abstract notion of influence and the more definite notion of STIT. We then consider modal agency operators and introduce an *other agent* extension which allows us to examine how these operators deal with cases of agents operating in parallel or in series. after discussing a number of operators we investigate them in terms of observed agent influence with our other agent extension. This allows us to describe cases where influence and strict STIT are similar and indicates what sort of reasoning we apply to manipulating influence operators.

1.1.4 Experimenting with our theory of influence

Our experimental work is in two stages, we have alluded to the presence of coaching agents and consideration of these agents is the first stage. Coaching agents are a bridge between theory and practical application. We discuss how coaching agents will use observations to generate hypotheses about agent influence and how these hypotheses may be ranked based on observed evidence. One of the major differences between theory and application is that theory admits unbounded concepts – the relentless forwards branching in a branching time model – and agents are bounded. We discuss this and propose a hypothesis structure that admits loops allowing a state based branching structure which provides a structure for linking hypotheses in a coaching agent's database of observations.

This database structure is then applied in a simple agent world and we experiment with its ability to detect agent influence in single agent and two agent settings in both noise free and noisy environments. The environment is then extended so as to allow agents to move around their world and we present results that indicate that an influence based approach is capable of identifying components of complex behaviour in noisy environments.

1.1.5 Influence and ability

Agents are bounded entities and this boundedness forces limitations on the abilities of an individual agent. These limitations may mean that an agent can only manipulate a very small part of its environment. Agents are also social entities, by working with other agents they may overcome their individual limitations.

To understand our notion of influence one should think of ability as being either *transparent* or *opaque*. A transparent ability is one that is fully under the control of agent, an ability that is fully contained by a single agent entity. Opaque ability is one which has a dependence on other factors, these factors may be other agents or tokens which extend an individual agent's abilities and which may not be immediately apparent to that agent or to a casual observer. Environment factors may be random, cyclic or episodic but they do not involve any element of *a priori* agent choice, our consideration is confined to other factors that are a result of agent choice.

Two agents simultaneously moving a large object that would be impossible for either to move individually; an agent handing a tool or token to another agent or an agent leaving a tool or token in the environment so that, at some future time, another agent may pick it up and use it. Each of these examples clearly involves another agent's choice. The simultaneous action in the first case lies on the boundary between transparent and opaque influence, if both agents know that jointly they can move the object and the agents are aware of each other's presence then the opacity lies in uncertainty about the other agent's choice. Given that agents are able to perceive their immediate environment this opacity may be dealt with by some signal or language where agents may indicate their intentions. Sequential actions are more complex, in the second example one agent may signal its intention to pass a tool or token to the other agent but this is not possible in the final example where the agents are never collocated. It is in these latter examples that we are most interested, the examples where one agent's behaviour may extend the potential influence of another agent. Where agents share an environment and are constrained or motivated by a common set of norms one could say that the first agent plays a part in *seeing to it* that the second agent carries out a particular action or brings about a particular environment state.

The notion of seeing to it is formalised as a modal logic in STIT theory. STIT theory has been extensively researched but there are problems in its abilities to represent nested other agent constructs, that is constructs representing one agent seeing to it that another agent brings something about. There are a number of approaches to managing such constructs but they carry heavy requirements for agents and agent societies.

We apply a novel notion of agent influence, rather than one of absolute agent ability, to STIT theory. In doing so we offer a viable reading for nested other agent constructs, a reading that carries very little in the way of additional society or agent requirements. This makes it potentially suitable for societies of reactive agents with limited cognitive abilities and fixed sets of actions. Applying our theory to reactive agents allows us also to explore some aspects of emergent behaviour, a potentially powerful but little understood property of societies of reactive agents. We make no claims for new approaches to reactive behaviour but do believe that our notion of influence may be used as a catalyst making the exploration of emergent behaviour in large state spaces a viable prospect.

Our notion of influence as one of opaque ability differs from other research on agent influence. It is, however, an entirely appropriate description and one which shall use throughout this work.

1.2 The contributions of this work

The main results of this thesis are the introduction of an interpretation of STIT which is based on the notion of agent influence removing objections to other agent nesting and the partial logical characterisation of this reading which allows for its implementation. These and the other contributions in this work are listed below:

A notion of influence which is offered as a reading for stit expressions in certain circumstances, most notably where individual agents play a part in a joint action or sequence of actions whilst maintaining the independence that agency confers.

A notion of *extended* influence which allows the representation of domains of agent influence characterising single agent influence, two agent influence and so on.

A notion of gateway actions which allow an agent or enables other agents to operate in a two or more agent influence domain.

A notion of strict stit which allows for the description of influence using the same terms as standard stit and which captures the differences in the way that values of propositions may be assigned to histories whilst satisfying the "sees to it" aspect of a statement.

Modal *leads to* and *may lead to* operators which pave the way for the treatment of the notion of influence in a similar manner to standard stit expressions.

A partial logical characterisation of *leads to* and *may lead to* operators which illustrates the differences between them and their similarities to standard necessity and possibility operators.

A binary representation of agent choice which simplifies the treatment of hypotheses that an action by an agent has a particular effect and allows for a simple treatment of refraining from said action whilst admitting the possibility that a different action by the agent may have the same effect.

The application of discrete time to the standard branching time framework which draws branching time into the domain of computation and provides a foundation for addressing the difficulties presented by *busy chooser* agents.

This list is summarised in the closing sections of this work where the the function of each contribution is outlined.

Chapter 2

Literature survey

We intend to describe a society where coaching agents observe the behaviour of simple actor agents and – by applying a theory of influence to these observations – tune the actor agent's behaviour so as to maximise their influence on their environment. This chapter is an introduction and brief survey of the extant research disciplines that provide the underpinnings of this system. Agents are the central theme of this work and are, consequently, the first area to be explored. After agents the role of beliefs and approaches to belief management are introduced then we consider the "system" aspects. It is intended that systems be specified by sets of safety and liveness properties, these sit, generally, under the umbrella of norms and normative systems. The final area is automated learning, the coaching agents are part of our agent society and their attempts to modify actor agent behaviour allow the society to be cast as a self contained learning system. Learning about one's environment is not a trivial task and learning to satisfy a set of rules or guidelines in a poorly understood environment is fraught with problems. This is an attempt at addressing some of these problems by applying extant techniques guided by a novel agent influence driven approach. Our contribution to the literature is the analysis of agent behaviour in terms of influence and the application of this influence to STIT logics so as to give an account other agent nested ability.

This chapter provides an informal introduction to the disciplines and research areas drawn together in this work. The chapter is broadly divided into two sections, the first introduces the areas that this work draws

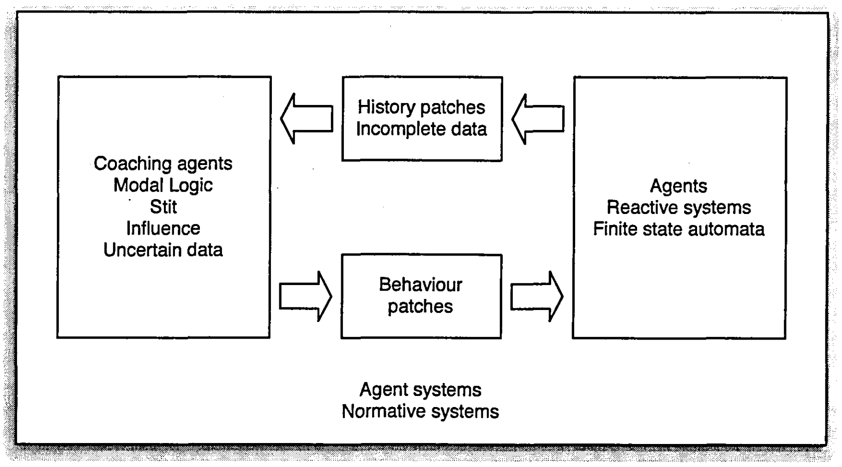


Figure 2.1: Schematic layout of a coached system of reactive agents

from and the second indicates, informally, how they may be drawn together so as to construct our society of influence driven agents.

2.1 A system outline: agents, agent systems, belief, logic and learning

We present a simple image of the system which will be a vehicle for this research. It is a multi agent system which uses reactive agents as primary actors. These reactive agents are monitored – by means of their leaving what we call "history patches" in the environment – by coaching agents which synthesise new behaviours for actors and leave these as "behaviour patches". Coaching agents aim to maximise agent influence in such a way as to encourage cooperation between agents. They aim to maximise agent influence whilst keeping the world within bounds imposed by system level norms, safety and liveness properties.

Figure 2.1 illustrates a simple system schematic diagram. The coaching agents and actor agents form a loop and communicate by way of data patches. Coaching agents may aggregate patches from a number of actor sources giving them a "broader view" of agent actions and abilities. These aggregated data are analysed for evidence of agent influence. The broader view afforded by data aggregation will be incorporated into the behaviour patches, which can potentially alter actor behaviour, and complete the agent / coach loop. It is worth noting that although we have mentioned agent / coach communications here communications proper play only a small role in the work. Our approach uses passive communications, coaching agents observe actor agents and actor agents may collect behaviour data synthesised by coaching agents - there is no real agent to agent communications path.

2.2 Agents, agency and agent systems

The concept of an agent has had a long history in philosophy[1]. These early notions were predicated on a *human* agent's ability to make decisions on the basis of his or her intellect. The computer centric notion of an agent entertained by this work is considerably more recent and seems to have grown from Reid G. Smith's work on the Contract net protocol (see Rosenschein and Zlotkin [97]). Research at this early stage was mainly concerned with ways of getting multiple software entities to interact appropriately and had no real concept of agent societies. Although separate software entities were involved there was still evidence of a centralised behaviour management or design. This is most notable where utility is considered at system level or where software entities have been designed with particular types of co-operation in mind. In time questions of self motivated entities, entities that considered utility on a *personal* rather than a system level, were raised. These did not fit neatly into the existing co-operation framework. Dealing with utility at a personal level gave these entities power to decide whether or not to co-operate with other entities. This early *distributed artificial intelligence* (DAI) (see Rosenschein and Zlotkin [97, page 14]), where specialised entities dealt with specific parts of a problem, grew into *co-operative problem solving* (CPS) (see Rosenschein and Zlotkin [97, page 15]) and agent research was launched on its own thread. There is still no notion of agent society and consideration, at this point, is focused on the self motivation of independent entities.

[1]See, for example, work by St. Thomas Aquinas (1225–1274 who wrote of human agents acting knowingly and willingly. John Duns Scotus (1265/66–1308) considered agents intellects as being more limited than Aquinas had considered them - possibly preempting the notion of reactive agents.

2.2.1 What does it need before we can call it an agent?

Co-operative problem solving's introduction of the notions of *personal utility* and *personal autonomy*[2] provide convenient starting points for an informal list of agent properties. A number of these definitions may appear to be rather general and almost seem to anthropomorphise agents and their properties. Indeed, Dijkstra (in a private communication now available at the University of Texas [38]) has warned against such anthropomorphisation. At this point these are best thought of as "convenience definitions" which make it easier to say what does not constitute an agent rather than what does. There is nothing contentious in any of these definitions and as this work progresses it will focus on specific types of agent removing the vagueness of anthropomorphisation. We proceed by outlining a series of working definitions of agent properties. From the personal utility and personal autonomy mentioned above we may observe that:

Observation 1 *Agents are self interested and self motivated.*

Self interest brings with it an implication of autonomy. Consequently an agent must have sufficient autonomy for it to be able to choose its actions independently. Autonomy is closely related to many aspects of agent behaviour and will be revisited below. For the time being we'll simply state that:

Observation 2 *Agents are autonomous.*

Autonomy and self interest are aspects not normally associated with what may be called *traditional* software entities. Building complex monolithic software systems is a difficult process. Complex interactions between modules may lead to unexpected system behaviour. Heterogeneous systems may multiply these difficulties and building heterogeneous systems of autonomous agents is a fearsome prospect. Agents may appear to act in similar ways but in a heterogeneous assembly of agents it is very likely that their internal machinations, the processes which cause an agent to act in a particular manner may be entirely different from those causing another agent to act in a similar manner. Comparing agents at the internal level makes no sense and, as mentioned above, it has become common practice to attribute agents with anthropomorphic "attitudes" such as beliefs, desires and intentions. The role of these attitudes when dealing with agents has long been recognised, see Bratman [18], Georgeff and Ingrand [47], and Rao and Georgeff [95]. This

[2]By *personal autonomy* we mean that an agent is free to decide whether or not to co-operate with another agent.

account of agent operation, abstracted by applying very "humanistic" descriptors provides a convenient and intuitive means of comparing the operation or performance of dissimilar agents and may be applied to even the simplest of reactive systems.

Observation 3 *Agents have desires and intentions.*

Desires and intentions imply an element of persistence. An agent may hold a desire to bring about a certain world state. If the agent does so and world state later changes then the agent will, once again, seek to bring that state about. If it is unable to do so, perhaps because it is no longer able, then it may either seek the assistance of another agent or to renew its ability. The (perhaps temporary) inability to achieve an intention does not necessarily mean that the agent will simply drop that intention, it will persist and the agent may find alternative means of achieving it. If the agent can't do anything about the problem then the only course of action available may be simply waiting. It is worth noting here that persistence, particularly with purely virtual agents, does not imply a continuous existence. If a purely virtual agent saves its internal state before its process is halted then it will be able to restart in the same internal state with its beliefs, desires and intentions intact. It may be, however, that it is reinstated into a different environment but persistence of intentions will mean that it may attempt to find other ways to operate. It is worth jumping slightly ahead to note that abstract concepts like intentions may be handled by modal logics, discussed later, allowing us to manage these anthropomorphic concepts in a formal manner.

Observation 4 *Agents have persistence.*

To justify being described as *self* motivated an agent must be able to generate and fulfil any intentions that it sees necessary as a means of achieving its desires through its actions. In order to make the most of its manipulations of its environment an agent must be able to perceive at least some of its surroundings. Generally agents are situated in their environment but, as with persistence, its worth noting that this "situation" need not be physical. A computer system remotely sensing and manipulating its environment by way of embedded sensors and actuators is, to all intents and purposes, situated in that environment.

Observation 5 *Agents are situated in an environment that they can perceive and influence.*

This rather informal description outlines the main features of agents in a very general manner. As mentioned above there is nothing contentious in these features and the intention is to provide a general description which will be refined as necessary later in this work. The final property – situatedness – implied that an environment was necessary and this is indeed the case. We continue by describing the properties of agent environments in a similarly informal manner. Environment details specific to our system will be discussed in greater detail in chapter 6.

2.2.2 The agent environment

Agents carry an element of situatedness, they need an environment to perceive and manipulate. Despite the lack of consensus (see Wooldridge [121] [122]) on what constitutes an agent it is safe to say that an environment is necessary – agents are nothing without an environment. Environment properties can be broadly divided into two categories, physical and societal. Here we address the environment's physical characteristics which we will see may support agent societies and societal behaviour. We consider these societal aspects in a little more detail in the following section.

Being able to hear confers no great advantages to agents which are required to operate in a vacuum. Being able to play soccer is not a great ability in a world where baseball is the only sport. The environment provides a collection of objects that agents (situated in that environment) can perceive and manipulate. This collection of objects also includes other agents sharing the environment, this means that any approach used to model agents should also be able to model their environment. Russell and Norvig [101] list five principal characteristics which may be used to describe and distinguish environments, these are uncontentious and are briefly described here with a view to indicating how these may require special attention in any model of an agent environment. Note that the descriptions here are from the perspective of a single agent situated in the environment under consideration. We draw heavily on descriptions from games here for the simple reason that their contrived and simple environments provide perfect examples of isolated environment characteristics.

- Determinism : If an agent is operating in a fully deterministic environment then the next environment state may be determined entirely by the present environment state and any action or actions carried out by an agent or number of agents. The results of agent actions are, therefore, predictable and consistent.

- Accessibility : Agents may not be able to perceive or manipulate certain elements of the environment. Games provide good examples; chess gives competitors a fully accessible environment where each player has perfect information on the state of the game before them. Games such as dominoes are partially accessible, a player may see his or her own pieces and what has already been played but is unable to see what any opponents are holding. Players do, however, have some knowledge of what other players are *not* holding. If an agent's percepts are limited to a subset of what its environment presents then this may have the effect of making a deterministic environment appear to be non deterministic. An agent that is unable to perceive rain may not be able to understand why the ground occasionally turns to mud and makes moving around more difficult.

- Episodic or non episodic : Episodic environments operate in cycles with each cycle being a completely independent operational entity, an "episode". In an episodic environment a poor decision in one cycle will not necessarily have any long term effect on the agents future utility. Games, once again, provide a good illustration; the card game snap is episodic, a moment's inattention may cause a player to lose one round but will not necessarily affect that player's chance of winning the next round. Chess is a non episodic game, a poor move in the early stages of a game may have serious consequences for that player's chances of making good moves in the future.

- Static or dynamic : The state of a static environment will not change during an agent's deliberation. A situated agent or agents will be the only drivers of change within that environment. If there is more than one agent then in order to maintain the static nature agents must co-ordinate their actions. Chess, again, provides a simple example with players taking turns to move pieces. Dynamic environments do not give agents the luxury of a change free period for deliberation. Elements of the environment may be in continuous change and this may be driven by uncoordinated agent action or natural events in that environment. Games, once again, provide several examples; chess is static because the players make alternate moves. Chess with a clock is semi static, if a player can complete deliberation and make a move within the specified time slot then the game is static, if not then the state game may change possibly rendering that agent's deliberations useless. Soccer is a dynamic game as the ball and members of both teams are moving continuously. Outside of games truly static environments are rare, Bratman [18] wryly notes that even if an environment contains no other human or robot agents then nature often intrudes.

- Discrete or continuous : if an agent's percepts and actions are limited, distinct and clearly defined then, as far as it is concerned, it operates in a discrete environment. In such discrete environments it is often easy to partition the agent's world into distinct states and easy to gauge the effects of an agents actions. Games, again, provide good examples, chess is a discrete game, each piece occupies one square on the board and can move only according to certain rules. Soccer is continuous, there are still rules governing movement but within these players can move freely in any direction and the ball can be anywhere in the field of play. Because of its discrete nature chess is amenable to brute force searches of the game state space by computerised systems, a technique that is signally unsuitable for soccer playing robots.

These five elements may be perceived differently by agents (and, possibly, differently by each agent) situated in an environment. Events that an external observer may see as being linked by an obvious, simple causal conjunction may not be obvious to agents situated in that environment. Any method used to model an agent system should be flexible enough to be capable of representing the system from the point of view of an observer as well as from the point of view of the least capable agent within the system.

2.2.3 Agent societies

We have a number of definitions related to agents and a number of definitions related to agent environments. This work is concerned with agent systems, as with agents there is no widely accepted definition.

We note that the ESSLLI course reader (Broersen et al. [20]) defines an agent system purely in terms of agents:

> There are many definitions for what a multi-agent system is (see Wooldridge [121]). For the purpose of this course, we define a multi-agent system as a set of acting and interacting, deliberating and communicating, autonomous, goal directed and socially engaged computing components.

This is, we feel, insufficient as agents are closely related to their environment and we have stated above that an environment is necessary. We therefore define an agent system as an environment with at least one agent. Because we are interested in the realisation of such systems in a computational setting we impose bounds on the environment. The nature of these bounds is unspecified but the import is that the environment consists of a countably finite number of elements.

Definition 1 *An agent system is a bounded environment containing at least one agent which is capable of perceiving and manipulating that environment.*

This definition covers both single agent and multi agent systems and is general enough to cover simple systems and complex societies where agents interact with each other as well as with their environment.

2.2.4 Agent communications

Our aim is to build a society of reactive agents which is capable of adapting its behaviour so as to satisfy system norms. We intend to have these agents co-operate so as to deliver complex, aggregate behaviours. What level of agent communication is required for this to happen? Bryson [27] suggests that there may be an imbalance in agent systems research with too much effort being put into designing complex communications systems for relatively simple agents. Similarly, Singh [108] notes that communications are centred on individual agents rather than on societal aspects of groups of agents. Even though humans have a rich communications language much is still accomplished by simple signals. Reactive agents are not ideally suited to using a complex communications language so we intend to use implicit communications and briefly outline this later in section 4.2.

2.2.5 Agent ability – agency

Wooldridge [122] uses the terms *agent* and *agency* interchangeably. This may be justified if the agents under consideration are *intelligent agents*, agents that act with a view towards achieving short term intentions and longer term desires. They are agents and they express their agency through their actions. Agents are not necessarily intelligent, Minsky [88] cites the example of a vehicle steering wheel. A steering wheel clearly has no intelligence yet the vehicle's driver may consider it as an agency for changing the vehicles direction. The steering wheel *appears* to know what to do in response to the drivers commands and the driver has no real need to know what the steering wheel does on his or her behalf. It is better to think of the steering wheel as an unintelligent agent that simply performs a given task. Consider, also, a coaching agent in a soccer team; without an understanding of what the game requires it would be very difficult for that coach to act deliberately so as to improve team performance. A soccer coach may be the oldest and least physically able member of a team. The coach may be unable to score the goals that the team require but the coach's knowledge and communication abilities in combination with the other team members physical abilities bring goals within their collective agency.

Clearly an agency is not the same thing as an agent and we state:

Definition 2 *An agency is an ability to carry out a particular task or bring about a required state. Such agency may be possessed entirely by an agent or jointly by a group of agents.*

Agents and agency are separate entities, the agent being a physical or virtual entity and agency being that entity's capability. Were they both to be considered as entities then the former would be a physical or virtual entity and the latter a metaphysical entity.

2.2.6 Representing agency

To work with the agency aspect of agents we need some means of representing the results of an agent's actions. Ideally this form would be portable across human and computer platforms. That is to say that the approach be computationally viable allowing computer based agents to use it and that it be to a sufficiently abstracted to allowing humans investigating system behaviour to understand its processes. Agency may be thought of as the behaviour of agents and this brings us, once again, to the touchstone phrase from Milner [87] that *the behaviour of a system is exactly what is observable.* This allows us to take an abstract view of the agents in a system and think of them as entities which behave by exercising autonomous choices.

2.2.7 Agents and agent systems, the intentional stance

Schut et al. [103] note that the idea of using humanistic concepts, concepts such as beliefs, desires and intentions, in the context of agents originated in work by Bratman [18] and Rao and Georgeff [95]. Although such abstract humanistic concepts imply that agents have a degree of cognitive ability they are still useful when dealing with reactive agents. Modal logics are capable of managing such concepts and these allow us to represent elements of these so called *pro-attitudes* (we consider this term in more detail in section 2.3.2) in a formal, non abstract manner. Bratman's approach identifies two problems that must be addressed by an agent's architecture; firstly the architecture must be able to support two forms of reasoning and secondly, the architecture must be able to deal with the problems of resource boundedness. The two forms of reasoning proposed are a means-ends reasoning approach to planning and an ability to reason about competing alternative courses of action. Resource boundedness can limit an agents abilities in a number of ways, limitations may be brought about either by the agent being unable to perform arbitrarily large computations or by the agent's environment continually changing.

The chess with a clock example, of section 2.2.2, is a good example of resource boundedness, agents playing the game have only a limited time to consider each move and may be forced into making choices that are not optimal.

Rao and Georgeff [95] described a computational model of a generic belief, desire and intention driven (BDI) agent. This generic agent adopted plans (which were precompiled procedures) as intentions dependent on environmental conditions, in doing so the agent may need to select one from several competing plans.

Wooldridge [121] notes that adopting an intentional stance by attributing beliefs, desires and intentions to agents has become common practice in agent systems research and notes elsewhere (Wooldridge [122]) that doing so ascribes *a strong notion of agency*. Considering agents as intentional entities provides a number of advantages; in a heterogeneous system these attributes can be a useful abstraction allowing researchers or developers to reasoning about agents without having to delve into their internal workings. Shoham [105] suggests considering artificial agents as formal versions of human agents holding formal versions of knowledge, belief, abilities and choices and indicates that this view helps with agent oriented programming. Shoham [106] also suggests that multi agent systems provide a number of intriguing possibilities for applying *mental attitudes* to assist in achieving coherent interaction between agents designed by different people. An intentional stance is useful when it helps understand the structure of a machine, its past behaviour or its future behaviour. There is a complexity floor below which adopting an intentional stance provides no advantage. An on/off switch could be thought of as a very simple agent, the simplicity of its behaviour means that there is nothing to be gained by adopting an intentional stance. We explore the notion of BDI agents further in the following section outlining agent architectures.

2.3 Agent architecture, cognitive and reactive agents

When faced with a new agent systems problem there is an immediate temptation to design a new architecture, a bespoke architecture perfectly suited to the task in hand. Wooldridge [121, page 235], however, notes:

> **You decide you want your own agent architecture.** Agent architectures are essentially templates for building agents. When first attempting an agent project, there is a great temptation to imagine that no existing agent Architecture meets the requirements of your problem, and it is therefore necessary to design one from first principles. But designing an agent Architecture from scratch in this way is often a mistake: my recommendation is therefore to study the various architectures described in the literature, and either licence one or else implement an 'off-the-shelf' design.

There are pitfalls in working with agent systems. This is a relatively young research area so there is a temptation, sometimes a great temptation, to try something completely new. There are many agent architectures and describing even a small number of these would take considerable space. Instead we simply partition the world of agent types into two broad categories, those of *cognitive agents* and *reactive agents*. Reactive agents, in their most basic incarnation, simply respond to cues, triggers and stimuli in their environment and do so by executing pre-programmed reactive plans – set sequences of actions – aimed at achieving particular goals. Jennings [68] notes that a major selling point of purely reactive agent systems is that overall behaviour emerges from interactions between component behaviours. This is view championed by Rodney Brooks [24], a strong critic of symbolic approaches to agency. The conceptual simplicity of reactive agents masks a number of difficulties, notably those of designing agents in such a way that they can take account of non local information and in such a way as to be able to improve their individual performance over time. Jennings further notes that agents using a large number of behaviours can quickly become too complex to understand.

Cognitive agents differ from reactive agents in a number of ways. The most notable being the cognitive agent's ability to plan by reasoning about its current state and its capability of maintaining models of its environment. These abilities intuitively carry advantages, an agent that can reason about things may develop new methods of carrying out tasks in a dynamic environment, its memory allows it to deal with non local perceptions and its model of the environment allows it to reason about the effects of actions.

Even though this work is concerned with reactive agents we have already indicated, following observation 3, that the notion of desires and intentions is not unique to cognitive agents. Despite their being only a small component of reactive agency we give this subject and considerable coverage so as to indicate the complexity required in explicitly managing beliefs, desires and intentions. This is a complexity that is not evident in reactive agents but in our coached system we need to be aware of this approach so as to allow us to abstract coaching agent operation where necessary.

2.3.1 The beliefs, desires and intentions architecture

We have already stated that the anthropomorphic notions of beliefs, desires and intentions are a useful abstraction for describing agents. This now common abstraction which describes a number of cognitive agent architectures grew from work by Rao and Georgeff following work by Bratman (see Trân et al. [114]).

Much philosophical literature considers intentions as an aggregation of beliefs and desires, Bratman's work argues that intentions are distinct from beliefs and desires. Rao and Georgeff [95] note that Intentions are partial action plans which may (or may not) help an agent to achieve its goals. The BDI family of architectures gives primary importance to these agent *intentions*.

BDI architectures provide only a very general framework leaving agent designers free to decide how an agent uses these beliefs, desires and intentions. With this in mind, rather than explore the specifics of interactions between the architectural blocks that build an agent we consider beliefs, desires, intentions and their management in a general manner.

2.3.2 Beliefs, desires and intentions – pro-attitudes

Bratman [18] notes that desires and intentions are *pro-attitudes*, that is attitudes which have a motivational role in agent behaviour. (Belief and knowledge are considered as *information attitudes* [122].) Believing that turning on an air conditioner will make me feel cooler will not motivate me to turn it on. This belief in conjunction with a desire to feel cooler may produce an intention to turn the air conditioner on. Desires and intentions may interact, a short term intention may interfere with an agent's ability to achieve a longer term desire so the agent will choose to suppress the shorter term intention. Intentions are conduct controlling pro-attitudes whereas desires are potential influencers of action. An intention involves a commitment to action that is absent from a desire, Bratman terms this relation between intention and action the *volitional dimension* of commitment.

The simple examples above use a belief-desire model, the intention to carry out a particular action may be reduced to the belief that given a desire for A then the action will bring about A. This model has both descriptive and normative aspects, it attempts to structure a common-sense approach to action and it attempts to articulate a practical rationality. Is this simple model adequate? If one accepts the existence of a predominant desire then it is. Earlier intentions may cause an agent to reason about what it needs to do and cause it to generate new intentions giving commitment a *reasoning centred* dimension. A prominent desire to bring about A means that an agent desires this more than other options which it may deem incompatible.

If an intention to bring about A cannot be identified with a predominant desire then it does not admit either a volitional or reasoning centred commitment.

The tactic of reducing intentions to beliefs and desires is, it would appear, inadequate. Bratman [19] considers intentions in the context of a boundedly rational agent. Bounded rationality introduces a requirement for *practical* as opposed to omniscient reasoning. This leads to intentions being considered as partial plans which can play a role in future reasoning. Intentions are both inputs to and outputs from an agent's reasoning process. As inputs they can pose problems – does an agent have the abilities it needs to bring about A? They can also pose constraints – does an agent's desire to bring about A interfere with any other desires or intentions? Intention has two facets, one deals with intentional action and the other with co-coordinating plans, recognising these different aspects of intention causes us to consider intention as a distinct element of agency.

Our brief discussion of agent societies, section 2.2.3, mentioned agents as being *autonomous, goal directed and socially engaged computing components*. What, then, is a goal and how do goals fit in to the BDI abstraction? Cohen and Levesque (see [43, page 301]) do not offer a formal concept of a *goal*, instead they consider only the consequences of goals. This leads to intention being considered as a persistent goal. Goals, in turn, are a subset of an agent's desires and form the set of desires that the agent has intentions of fulfilling.

This brief discussion may be summarised by a number of definitions:

Definition 3 *A pro-attitude is something that plays a motivational role in an agent's behaviour.*

Definition 4 *A belief is a fact that an agent holds to be true in its present state and which may be true in the future.*

Definition 5 *A desire is a pro-attitude which can, potentially, influence an agent's behaviour.*

Definition 6 *An intention is a pro-attitude which controls an agent's behaviour and to which the agent has some degree of commitment.*

2.3.3 Belief management

Understanding how beliefs ought to be changed given new information has been an active research area in both philosophy and artificial intelligence. We cover this briefly and informally here, our work makes no explicit use of belief management but the underlying notions serve to guide our consideration of coaching agent operations. Two approaches to belief management have been studied in detail, *belief revision* and *belief update*. Agent beliefs may be described in two ways, a *belief set* is a set of beliefs that is closed under logical consequence. A *belief base* does not exhibit such closure and may, therefore, contain beliefs that are independent of others. Belief revision is concerned with the changes an agent should make to its belief set when a new belief is adopted. Friedman [45] notes that belief update addresses how an agent ought to change its beliefs when it perceives a change in its environment. An agent's beliefs will, to a large extent, concern its environment so there is a subtle link between these approaches. Both approaches share the intuition that changes to belief sets ought to be minimal. The changes made by each approach differ, belief revision will attempt to identify beliefs which ought to be discarded so as to accommodate a new belief whereas belief update attempts to identify what changes to existing beliefs are necessary to accommodate a new belief.

2.3.4 Belief revision

Belief revision originated from studies in the philosophy of science and revision occurs following the adoption of a new belief. If an agent learns φ and φ is consistent with its existing beliefs then the new belief, φ, is simply added to the agent's knowledge base. If, however, φ is inconsistent with existing beliefs then these are revised by discarding older beliefs so as to maintain consistency. Friedman and Halpern [45] The most commonly accepted approach to belief revision is known as the AGM theory after Alchourrón et al. [3] and Gärdenfors [46]. This assumes that the agent's epistemic state is represented by a set K of formulas in some logical language, $\mathcal{L}_e$ over a set of primitive propositions. Belief revision takes a set of beliefs, A, a revision operator $\circ$ and a new formula, φ and, after the operation, returns a new belief set, $A \circ \varphi$. Intuitively this process should result in a *minimal change* to the existing belief set. Friedman and Halpern [45] list a set of postulates to characterise this notion.

The AGM theory's assumption that an epistemic state be represented by some logical language requires a little consideration. Recall that our work is biased towards reactive agents which are not the natural habitat of such systems. To this end we describe beliefs as the agent's current vision of its world, a vision which is limited to its percepts.

Our agent's beliefs may, then, be characterised as being of the form *given a set of percepts P then action α is the best action.* The choice of the best action falls into the remit of coaching agents which will guide actor agents in their behaviour. Beliefs, albeit in a very simple form, do exist in tangible form in a coached reactive agent system.

2.3.5 Belief update

Belief update originated from work in the database community (see Friedman [45]) and addresses the problem of changing a knowledge base so as accommodate new facts or beliefs about the world. If an agent makes a new observation that contradicts or is inconsistent with existing beliefs then existing beliefs are not necessarily considered as being false. Belief revision does not assume that the world is unchanging and, instead, attempts to capture changes in the world. This raises some interesting questions as there are subtle differences between static and dynamic environments and differences between changing and unchanging environments. We consider static and dynamic as describing worlds where the *properties* of objects remain constant (see, for example, Russell and Norvig [101, page 46]) and the disposition of world objects may or may not change during agent deliberation. A *changing* environment is one where the properties of objects are subject to change as, for example, an agent learns more about them. An *unchanging* environment means that object properties do not change though this, clearly, may still be a dynamic environment.

Katsuno and Mendelzon [71] characterised the belief update procedure as a set of postulates addressing formulas, rather than belief sets, which an update operation ought to satisfy.

2.3.6 Some philosophical points

Belief management is a lively area of philosophical research and one that autonomous agent systems may contribute to. The inherent boundedness of agents allows the tuning of constraints so as to address certain issues, something which would not be possible with more complex human systems. Hansson [58] lists a number of philosophical concerns with belief revision, since some of these are pertinent to this area of work we briefly outline them here.

The AGM theory is unusually simple and elegant but this elegance comes with a cost. The simplicity is a result of assumptions about belief systems and AGM does not capture many of the subtle aspects of belief management. The simplicity and elegance is also rather fragile and much of it lost when postulates are added to make the model more realistic. Belief change processes may be extremely complex and are, consequently, subject to idealisation so as to produce tractable models. This idealisation is generally by one of two methods, simplification where complexities are left out, and perfection where standards of rationality are raised beyond what an agent may reasonably achieve. Most research has taken the latter approach resulting in simpler mathematical models which require unlimited cognitive capacity. Real agents may be bounded by many constraints making such an approach undesirable. Finiteness, although a weak restriction, takes us towards a realistic representation of cognitive capacity, Hansson asks if *stricter cognitive limitations than finiteness may be represented in an interesting way?* Belief management in a multi-agent environment may go some way towards answering this question. Agents are by definition bounded entities and may face additional constraints such as limited abilities, having to depend on assistance from other agents and relying on delegated tasks or duties.

The logic of the BDI model.

BDI components are usually considered against a background of branching time and possible worlds - a standard approach for models incorporating logical modalities. We have noted that modal logics are capable of handling such concepts and for some examples see Halpern [55], Horty [66] and Wooldridge [122]. Rao and Georgeff build on earlier work using possible worlds frameworks and consider intentions as being on par with beliefs and desires (this *elevation* of importance is in line with Bratman's consideration).

Possible worlds semantics traditionally treat each world as a set of propositions and model beliefs as a *belief accessibility relation* linking certain worlds. Halpern and Moses [53] note that a proposition is believed if and only if it is true in all belief accessible worlds. Each belief accessible world leads, in turn, to a tree of further worlds. Intentions can be represented, similarly, by a set of *intention accessible* worlds. Recall that an intention involves some degree of commitment, this allows us to characterise the set of intention accessible worlds as a set of worlds that the agent has committed itself to attempting to realise.

Rao and Georgeff [95] describe these in three accessibility relations, one for beliefs, one for desires (which may be applied to goals) and one for intentions. These relations deal with two *types* of attitude, informational and motivational. Bennett at al. [9] indicate that informational attitudes may be characterised by a KD45 modal system[3] and motivational attitudes by a KD system.

2.3.7 Dealing with beliefs, costs and benefits

Beliefs may or may not be true but knowledge should always hold. Truth maintenance is a well researched area of traditional artificial intelligence. A truth maintenance system (TMS) ensures that any reasoning currently in progress and based on earlier assumptions is updated when these assumptions are validated or invalidated. The truth maintenance system may be integrated into, for example, a diagnostic system in such a way as to make its operation implicit. When alarms or information messages arrive they are checked against existing hypotheses and, if necessary, a new hypothesis is instantiated. The hypothesis list is then evaluated

Recall that agents have both goals and intentions and in section 2.3.2 we sketched the relationship between goals and intentions. An agent should only maintain an intention for as long as it is feasible to do so, unnecessary intentions may carry penalties in terms of additional resource usage. Clearly an agent must have a policy for reconsidering its intentions.

[3]Modal systems are described in a later section of this chapter.

There is a body of related work which looks at managing the intentions of single agents. Schut and Wooldridge [102] examined costs and agent efficiency associated with reconsidering intentions in a complex environment. Earlier work on re-evaluating by Kinny and Georgeff [73] has shown that there is no best approach and that the approach used is dependent on the environment.

Kinny and Georgeff [73] deal with single agents, dealing with a task in a multi agent system may involve several agents co-operating and, perhaps, working to a group plan. Let us consider this plan as an embodiment of the intentions, or a meta intention, of that group of agents. Intentions are a future directed route map towards a goal. In a dynamic environment things may change unexpectedly, because of this the validity of intentions, in the context of achieving particular goals, may change. Any reconsideration of intention at an agent level may have knock on effects for other members of the group and may, possibly, render the whole plan invalid. One agent reconsidering its intentions may cause a violation of the intentions of another agent. In a dynamic environment intention violations are not only generated internally, other single agents or groups of agents working independently may change the environment in such a way as to violate one or more agent intentions in our plan.

2.3.8 Reactive agents

Cognitive agents, as their name implies, are capable of deliberation. Reactive agents, in their most basic incarnations, simply react to their environment but, importantly, Chang et al. [28] note that reactive systems engage in *constant interaction* with their environment. Intelligence and rationality appear in many different forms, humans are cognitive rational entities, entities that deliberate their actions and plan days or years in advance. Ants simply respond to their environment, although an individual ant may not exhibit intelligence in the same way that an individual human can it is hard to deny that ants, collectively, are rational.

In our system the actors are purely reactive, their percepts are interpreted as patterns which trigger the choice of a behaviour or set of behaviours. There is a rationality in these choices and that rationality seeks to maximise agent influence.

Ferber [43, page 16] indicates that one of the difficulties raised by reactive agents is that their mechanisms of reaction to events rarely admit to an explanation of agent goals or planning mechanisms.

2.3.9 A reactive bias

We have already stated that this system will be based on reactive agents and any following consideration of cognitive agents should be read with this in mind. Although the simpler reactive agents have no explicit belief management the fact that our proposed system will learn by adapting its behaviour indicates there something of this form is happening somewhere in the system. The consideration of belief management is here as a means of providing a context for our reactive agent's coached behaviour adaptation.

We make some passing mention of hybrid agents here, these combine aspects of reactive and cognitive agent architectures and attempt to pick the best of both types. There is a body of literature on hybrid architectures covering areas such as process management agents, Debenham [37], hybrid control of robots using an interface agent, Strippgen and Peters [109], and symbolic/reactive hybrid robot control, Oliveira [89].

2.4 Normative systems

Our intention is to have a system of reactive agents develop appropriate behaviours to satisfy some system specification. We need, at some point, to say what is and is not appropriate behaviour and we do this by treating the system – the coaching agents and the acting agents – as a *normative system* and that is simply a system controlled or regulated by rules. These rules, or *norms*, either require agents to behave in a particular manner or forbid them from certain activities. A speed limit is a good example, a limit requires agents to drive at or below a stated speed and forbids them from driving at speeds greater than the stated limit. Agents are free to decide whether or not to violate the speed limiting norm, if they do then they risk sanction by way of suspension of privilege, loss of utility or both. One of the aspects of a normative system that makes it especially suitable for our approach is that norms *can* be violated. If an agent brings about a bad world state then that agent will not be individually sanctioned but the behaviour that led to the bad state will.

The issue of norm compliance brings a number of problems. Ågotnes et al. [2] discuss this issue with a view to logically describing the robustness of a normative system. For the time being we are concerned with the synthesis of behaviours which will satisfy norms and simply accept that there will be some violation as part of this process. Future work will investigate robustness and system recovery from norm violation.

2.4.1 Safety and liveness

A common systems representation of norms is safety and liveness properties. Safety and liveness properties were introduced by Lamport [78] and state, intuitively and respectively, that something bad will not happen and that something good will eventually happen. Lamport's intention for safety and liveness properties was the verification of the correctness of multiprocess programs. A crucial aspect of safety properties is that the violation of a safety property may always be detected by a finite prefix or sequence of behaviour. We are interested in reactive agents and one of the characteristics of such agents is a constant interaction with their environment. We assume, therefore, that an agent's behaviour is infinite. This is not a restriction (see, for example, Kurki-Suonio [77, page 62]) as terminating executions (such as the agent bridging state in the second of the experiments detailed in chapter 7) may be represented as a behaviour where a terminal state is repeated indefinitely. If a safety property has been violated then, Kurki-Suonio [77, page 63] notes, it can not be remedied in the future. It may be the case that violation of a safety property may lead to a bad situation that can not be remedied but in our system agent actions are local and the fact that the violation can be detected means that coaching agents may take steps to prevent future occurrences of the violation.

This is the approach that we adopt for specifying agent behaviour.

2.4.2 Norms are social

Conte and Castelfranchi's book Cognitive and social action [33] argues that norms are typically a social phenomena and this supposition of a social aspect makes normative systems intrinsically multi agent systems. Certain norms from a given set may be for a certain class of agents, a subset of the system's population. For agent classes norms may guide an agent's desires or the assignment of tasks. Boella and van der Torre [13] indicate that normative systems share a number of properties with autonomous agents and, thus, may be considered as being agents in themselves. This implies that the normative system governing a society or system of autonomous agents may be considered as being some form of meta agent. The autonomy of each agent that is a *member* of the normative meta agent is preserved so the system remains a proper agent system. In a dynamic society the composition of this meta agent will also be dynamic, it is composed of certain interactions between real agents, physical or virtual and these interactions may be constantly changing. This leads to a question; what is it in a system that binds certain agents together to make the normative meta agent? We attempt to address this question rather crudely by embodying the normative agent and characterising it as a specialised coach. We are interested in the emergence of system norms so this coach is tasked with identifying the norms required to guide agent behaviour. In order to formalise norms so that they can be analysed and manipulated we adopt the approach suggested by Boella and van der Torre [14] and use deontic logic.

2.4.3 Normative architectures

Several normative agent architectures and systems have already been proposed. Boella [12], for example, outlines NMAS, normative multi agent systems based on a normative interpretation of Searle's [104] speech act theory and a *counts as* operator. Stütz and Onken [110] describe a normative system for pilot assistance that uses case based reasoning to modify a Petri net that acts as the normative in a system. Other examples include the BOID architecture, Bratman [19], and Kollingbaum and Norman's NoA architecture [75]. These systems focus on cognitive agents, agents that are able to reason about the consequences of norm violation. Such agents may deliberately violate norms if the perceived benefit outweighs the risk of sanction. We are interested in systems of reactive agents, agents that do not model their environment and are, consequently, unable to reason about possible sanctions resulting from norm violation. Boella and Damiano [11] describe a reactive architecture for an agent in a normative environment. Their reactive agent has a deliberative module which deals with planning and re-planning and they allow for deliberate norm violation. This is not the approach which we adopt as we are looking for norm compliance to be implicit in agent behaviour but accept violation as a part of the system's exploration and learning sequence. This could be certainly be achieved by simply designing behaviours that do not violate norms but such hard wired ideal behaviour would not allow us to take advantage of one of the benefits of normative systems – adaptability. Boella and van der Torre [15] indicate that an important feature of norms is that they allow for behaviour that deviates from ideal. Generally deviation from ideal behaviour may invite sanction and, in our case, sanctioning will be by coaching agents biasing actor agents against acting in that manner in the future. In normative systems agent autonomy has a higher priority than norms, the process of deciding whether or not a behaviour violates norms is an autonomous activity, as described by Boella and van der Torre [13], and one that is carried out by an observer in our system.

Boella and van der Torre [15] further note that norms may be considered as *soft* constraints in a system which detects and sanctions violations. *Hard* constraints, such as a token entry system, which make violations impossible are not really norms. This may be the case for social systems but recall that we wish to characterise system norms as safety and liveness properties. Safety properties are norms that *must not* (as opposed to *should not*) be violated but it would be difficult to have a system which learns by exploration never violate a norm. In order for a system to know how to observe norms it must know what violates norms.

One worrying aspect of sanctions is that they may be episodic in nature, as described by Russell and Norvig [101, page 46], and only affect an agent in one of a number of roles that that agent has, this may lead to the agent simply shrugging its virtual shoulders and not modifying its behaviour so as to avoid norm violation. We adopt the notion of a *bad state* or, more accurately, reduced utility as our implicit system level sanction. Utility, in our agent view, is influence and reduced influence is non episodic and, thus, more likely to be a useful behaviour modifier than an episodic sanction which may be taken out of context.

2.4.4 Managing norms, virtual agents and virtual societies

Boella et al. [12] describe a conceptual model of virtual organisations. This separates norms guiding agent behaviour into a separate normative system and then treats that normative system as a virtual agent. The normative system, then, defines roles that agents in the system may adopt. This allows the characterisation of agents and the normative system – which is modelled as an agent – as a recursive game. Although treating the normative system as another agent brings advantages, agent interaction with norms may be characterised in the same manner as interaction with other agents, it is unsuitable for our proposed system for a number of reasons. The prescriptive nature of the normative system does not fit well with our intention to partially address emergent behaviours, outlined in section 2.12. Emergent behaviour implies that there is no hierarchy of agents and that agents themselves adopt roles according to their abilities – there is no prescriptive role assignment. The normative systems described by Boella and Hulstijn [12] have a sanctioning system and one of the roles described is that of "defender" agents. Such defender agents may be useful in a setting where norms can be guaranteed system wide but is potentially dangerous in settings where an agent may travel from one locality to another where there are different norms. If an agent travels from one system to another then defenders sanctioning newcomers will help to maintain system stability and educate new agents. If a defender moves to another system then it will, most likely, be poorly equipped to dispense judgement and, perhaps fortunately, it will be unable to apply sanctions. If we apply a world view to our system than it may be possible for a defender to move to another part of the world where its beliefs of what is proper do not hold but it still has an ability to impose sanctions. This is clearly not good. If an agent moves to a neighbourhood where its norms are not applicable then, in our system, that agent learns directly by utility reduction rather than by sanctions which are, possibly, being misapplied. Unless norms can be guaranteed system wide – which implies a degree of centralisation – then the idea of defender agents is untenable. Norms in our system may be thought of on two levels, the observer has notions of what the system ought to achieve and coaching agents have notions of maximising influence. We believe that the use of coaching agents, agents which educate rather than sanction, is a rather better approach allowing a decentralised system without a hierarchy or agent roles.

2.4.5 Reactive agents in normative environments

Our aim is to have agents operate in as *pure* a reactive manner as possible. Such an operating mode will mean that norm compliance is implicit in an agent's behaviour. This requires that agents are equipped with a library of reactive behaviours that are sufficient to allow them to function effectively with minimal additional resources. Other approaches to reactive normative agents are encumbered by unnecessary support systems, our approach is designed so that an agent may shed such support when it is no longer necessary. Our approach builds on exploratory work by Logie et al. [82] where uncoached self organisation was investigated. Adding a coaching level abstracts the system's self organisation to a point where we may formally reason about it.

2.5 Learning systems

There are numerous approaches to machine learning and coached agency so before examining what is available it is, perhaps, best to say what we are and are not interested in. Our aim is to develop a system for *reactive* agents which allows them to develop behaviours suited to satisfying environment norms. The lack of a world model and cognitive abilities makes reactive learning seem like a difficult task. One potential approach is to embed a neural network in the agent's core. With suitable training or evaluation systems such an approach may well produce an agent that is suited to its environment and its associated norms but that agent and its behaviours are very tightly bound together. It is next to impossible to excise a single useful behaviour from a neural network which aggregates several behaviours. This is an important consideration, multi-agent systems may contain a number of similar agents or agents belonging to the same *agent-equivalence* class and we may wish to transfer behaviours developed by one agent to other agents.

2.5.1 Two broad categories of machine learning

Theories of machine learning fall into two broad categories [60], classical learning where a decision space is searched for a good fit to data and Bayesian learning which can be viewed as a process of reducing uncertainty. The former approach is not really suited to reactive agent environments as equipping agents or coaches with such abilities runs counter to the idea of simplicity. Coaching allows this by placing a coaching agent in a *behaviour synthesising* role. Learning may be a misnomer, coaching agents will, in reality, observe agent behaviour and synthesise new behaviours. Learning occurs when this is viewed as a system but no individual agents carry a complete learning process. The sections below very briefly outline some of the concepts involved in agent learning.

Russell argues that most AI learning research has focused on environments that are static, deterministic, discrete and fully observable [100] [101] as opposed to dynamic, stochastic, continuous and partially observable agent worlds that are of interest to us.

2.5.2 Reinforcement learning

Recently, reinforcement learning has made progress in agent applications. Shoham et al. [107] note that though reinforcement learning has been an active topic in general AI research there is still a very small body of research on multi agent reinforcement learning possibly because there is no clearly defined problem statement, Bowling [16], for example, notes that reinforcement learning research has tended to focus on single agents.

Reinforcement learning methods are aimed at solving fully observable Markov decision processes. Such processes consist of a reward function and a model, there is no real difficulty in accommodating the model of a reactive agent as it is simply a set of state transition probabilities. Reinforcement learning algorithms can be model free as in, for example, Q learning [111]. These approaches work well in fully observable worlds but, in agent systems, partial observability is rather more likely. Earlier work by Åström [5] [100] has proved that optimal decisions in partially observable Markov decision processes depend on the belief state of an agent. McCallum [85] has shown a way to approximate an agent's belief state using recent sequences of percepts.

2.5.3 Collective intelligence

In order to be able to implement some form of reinforcement learning agents must have some idea of states that are preferable over others and that preference must be ordered in such a way as to allow the agent to see utility as some form of reward. Wolpert et al. [120] consider the problem of automated design of large scale decentralised systems and use the term *CO*llective *IN*telligence – COIN – to describe systems which embody solutions to this problem. Wolpert's approach uses an agent level reward function that that is updated so as to drive agents towards achieving a global goal. One of Wolpert's main concerns is not having agents work at cross-purposes. However, this system has a function which considers utility at a "world" level which is something that we wish to avoid.

2.5.4 Evolutionary learning

Evolutionary systems are an active research area. Hoen and de Jong [64] note that multi-agent learning means that an agent must learn to select actions that maximise their utility given the action choices of other agents. Indeed, Hoen and Bohte's work [63] is close to what is being covered here but its evolutionary nature makes it rather different from this work's learning system. Evolutionary learning, like genetic algorithms, is population based and relies on producing generations of agents. The selection process, in common with genetic algorithms, applies some form of fitness function to a population in order to select candidates for carrying forwards to the next generation.

Our agents learn but they do not evolve, by this we mean that they start with a fixed set of percepts and abilities and these do not change during the agent's life. It may well be that an agent's behaviour, which in the case of reactive agents is a mapping from percepts to actions, changes so as to effectively suppress a particular action. This action will, however, remain with the agent and if its environment changes so that the suppressed action becomes necessary then it is available.

Hoen and de Jong's work is based on the COIN framework and raises questions of global utility, we avoid addressing anything at a global level. Further, Hoen and de Jong's approach attempts to evolve joint actions by decomposition, we adopt an approach which observes, searches and maximises influences without the need for considering utility and reward functions.

2.6 Modal logic

We have indicated that the notion of *seeing to it* is a useful and powerful tool for representing agent ability. Modal logics and the closely associated concept of STIT and our characterisation of this as *agent influence*.

Modal logics provide the formal core of this work and, consequently, we devote more to these concepts than other contributory areas. Here we introduce modal logic as a foundation for stit theory which we outline in the following section.

2.6.1 A brief, informal introduction

Modal logic concerns itself with the *modes* in which things may be true or false [93, page 20], in particular these modes are necessity and possibility. At its most simple, modal logic extends well formed *truth functional* propositional logic formulas by the addition of two modal operators, $\Box$ and $\Diamond$ representing necessity and possibility respectively. Before I leave my apartment in the morning I look at the sky and decide that rain is a distinct possibility, consequently there is a chance that I will get wet. This may be expressed in numerous ways given these modal operators: "it will possibly rain so there's a chance that I'll get wet" – $\Diamond$*Rain*$\rightarrow$ $\Diamond$*I will get wet*. If my wife tells me not to get wet then I may reason that "there's a chance that I may get wet so it may rain" – $\Diamond$*I will get wet*$\rightarrow$ $\Diamond$*Rain*. To ensure the possibility of my staying dry it's necessary for me to take an umbrella when I leave the house, $\Diamond$*I will stay dry*$\rightarrow$ $\Box$*I take an umbrella with me*. Modal logic is commonly referred to as a logic of *necessity* and *possibility* [30] [42] but Fagin [42] notes that these terms should not be taken literally as they are very context dependent. *Necessarily*, he states, could mean *according to the laws of physics* or *according to my beliefs*, this is a vital element which gives the flexibility to allow its use in heterogeneous environments.

2.6.2 Many types of modal

Modal logics are the result of propositional logic being augmented by operators for necessity and possibility. There are a number of modal logics which provide different readings of necessity and possibility. Deontic logic adds operators representing concepts such as *it is obligatory that...*, and *it is permitted that...*, and *it is forbidden that....* Temporal logic adds operators representing *it will always be the case that...*, *it will be the case that...*, *it has always been the case that...* and *it was the case that....* Doxastic logic adds an operator to represent the case where x believes that .... These are all members of the modal family. In this work we will consider basic modal logic and, later, examine some details of deontic logic.

2.6.3 Branching time, possible worlds and truth in modal systems

In 1962 Jaako Hintikka recognised that it was possible to characterise an agent's beliefs as a set of, what he termed, *epistemic alternatives* and proposed that they be represented by a *possible worlds* model. Saul Kripke later developed semantics allowing Hintikka's ideas to be formulated in terms of possible worlds in a modal logic. Arthur Prior introduced the theory of *branching time* [94] and this theory was developed further by Richmond Thomason [112] [113].

We begin building a branching time framework by considering here and now. Looking backwards in time there is a single history leading here and now, things may have turned out differently but here and now they are as they are and the thread of past events is settled. The future is different and offers many possibilities, some of these possibilities are the result of chance, some the result of individual decisions or actions and some the result of third party decisions and actions. Here and now is a moment where a single past divides into a number of futures. Since branching time is one of the core components of this work we begin listing definitions relating to it.

Definition 7 *A moment is where a single thread of events divides, it partitions time separating the single thread of settled past events from an indeterminate number of future threads. The division may be the result of agency or of chance.*

The notion of a moment is to capture an event which has some effect on how the future develops. A moment has an implicit sense of before and after and it is clearly silly to have all events occurring at the same time. We introduce a transitive and irreflexive ordering, $<$ on sets of moments. Because of the nature of the passage of time through these moments, a single settled past and multiple futures, any ordering will render a set of moments into a tree like structure. The ordering, $<$, imposes is a partial, tree-like ordering on a set of moments, $Tree$, such that:

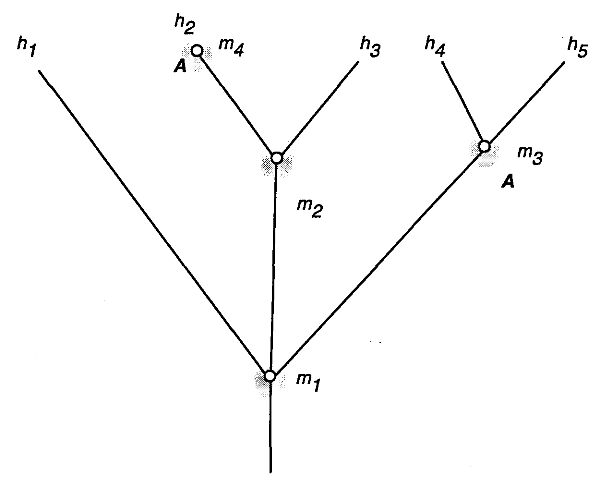


Figure 2.2: Histories and moments arranged in branching time

Definition 8 *For any three moments m_1, m_2 and m_3 from $Tree$, if $m_1 < m_3$ and $m_2 < m_3$ then either $m_1 = m_2$ or $m_1 < m_2$ or $m_2 < m_1$ (from [66]).*

This allows a structure similar to that of figure 2.2 which illustrates four moments, $\{m_1, m_2, m_3, m_4\}$ arranged as a tree. We list further definitions pertaining to the ordering of moments:

Definition 9 *A set of moments, M, from $Tree$ is said to be linearly ordered whenever, for any moments hm_1 and hm_2 belonging to M then either $m_1 = m_2$, $m_1 < m_2$ or $m_1 > m_2$ (from [66]).*

This linear ordering on a subset of moments from $Tree$ allows us to extract a single path through the tree. This single path represents one possible evolution of events from the earlier moments at the bottom of the tree to the later moments at the top and we refer to it as a *history*.

Definition 10 *A linearly ordered set of moments, h, is a maximal linearly ordered set when it can be no larger whilst remaining linearly ordered. Such a maximal linearly ordered set of moments from $Tree$ is called a history (from Horty [66]).*

It is occasionally useful to consider a number of moments occurring at the same time in a tree of histories. Belnap and Perloff [34] introduce the notion of an *instant* to do this.

Definition 11 *Given a branching time tree, Tree, containing a set of moments, M, and histories, H. An instant is a horizontal partitioning of the tree into equivalence classes such that each moment contained in an instant is said to occur at the same time.*

With this definition we can see that figure 2.2 represents a tree containing four moments arranged in five histories. Because of the indeterminism of time a moment may be contained in more than one history, m_2 is contained in both h_2 and h_3 We say that $H_m = \{h\colon m \in h\}$ represents the set of histories passing through moment m, the set of histories in which m occurs. Combining $Tree$ with a transitive, irreflexive ordering, $<$, on its members gives a *branching time frame*.

Definition 12 *A branching time frame is a structure $\mathcal{F}$ of the form $\langle Tree, <\rangle$ with $Tree$ a non-empty set of moments and $<$ a transitive, irreflexive ordering on $Tree$ (from [66]).*

We now turn to the evaluation of the truth of a formula against this background of branching time. It is straightforward to evaluate the truth of a formula at a given moment. The fixed nature of the history leading *to* a moment makes the evaluation of a past operator, P similarly straightforward. Evaluating the truth at future moments by, say, a future operator F, is not so easy. In figure 2.2 we see that A is true at moment m_4 but what can we say about the future truth of A, $\mathsf{F}A$ at the earlier moment m_1?

Some approaches, notably that of Prior and Thomason, are unable to address this question. $\mathsf{F}A$ is true at m_1 with A really being in the future if either h_2, h_4 or h_5 is achieved and false if h_1 or h_3 are achieved. Each of these is possible at m_1 and this limits what we can say. A moment alone, it would appear, does not provide sufficient information to allow evaluation of $\mathsf{F}A$. Prior and Thomason suggested that the evaluation of future directed statements must be against not only a moment but also a history passing through that moment. This index can be represented as m/h, a moment m and some history $h \in H_m$.

If this approach to evaluating future directed statements is adopted then for the sake of semantic consistency other formulas ought to be evaluated in the same manner. Branching time models thus become structures of the form $\mathcal{M} = \langle \mathcal{F}, v\rangle$ where $\mathcal{F}$ is a branching time frame and v is a valuation function which maps each propositional element onto the set of m/h pairs where the propositional element in question is true. The basic truth definition for branching time models indicates that propositional elements are true where v says so. We list definitions for evaluation rules for basic operators.

Definition 13 *For an atomic formula, A, a valuation function v and an index m/h from a branching time model $\mathcal{M}$:*

$$\mathcal{M}, m/h \models A \ \textit{iff} \ m/h \in v(A) \tag{2.1}$$

$$\mathcal{M}, m/h \models A \wedge B \ \textit{iff} \ \mathcal{M}, m/h \models A \ \textit{and} \ \mathcal{M}, m/h \models B \tag{2.2}$$

$$\mathcal{M}, m/h \models \neg A \ \textit{iff} \ \mathcal{M}, m/h \not\models A \tag{2.3}$$

$$\mathcal{M}, m/h \models PA \ \textit{iff} \ \exists\, m' \in h \ \textit{such that} \ m' < m \ \textit{and} \ \mathcal{M}, m'/h \models A \tag{2.4}$$

$$\mathcal{M}, m/h \models FA \ \textit{iff} \ \exists\, m' \in h \ \textit{suchthat} \ m < m' \ \textit{and} \ \mathcal{M}, m'/h \models A \tag{2.5}$$

(From [66].)

More formally, Kripke semantics are usually represented as structures and the most usual is a triple [59] $\mathcal{M} = (W, R, P)$ where W is a set of all possible worlds, R is a binary *accessibility relation on* W and P is a mapping from the set of natural numbers to subsets of W.

How does this relate to an individual agent's situation? Recall that an agent is a bounded entity situated in an environment and that an agents sets of percepts and actions are finite. At a given world an agent may perceive certain things and use these with its knowledge and available actions to generate a number of possible future worlds. Intuitively R encapsulates an agents knowledge and abilities in specifying which worlds it may consider as being *accessible* (or *possible*) from its current world. This leads to informal definitions of belief and knowledge. An agent *believes* a proposition if that proposition is true in at least one future world and an agent knows a proposition if that proposition is true in all future worlds. Knowledge and beliefs are precipitated on actions, recall that definition 1 stated that an agent can at least partially manipulate its environment. This also addresses environmental determinism, if an agent is certain of the consequences of an action then it knows that it is able to see to it that something is brought about, most likely as a result of an action or sequence of actions by that agent. Horty [66] covers this aspect thoroughly and uses it to express ability predicated on an agents choice of action, this idea will be explored in detail in the following sections.

Given that an agents beliefs about future worlds are predicated on actions (note that an action may be to refrain from doing something) we may now define the notions of knowledge and belief predicated on an agents actions. Belief and knowledge carry implications of cognitive ability but the notions are applicable, at an abstract level, to reactive agents. Recall that our coaching agents operate by observing agent behaviour and generating hypotheses based on these observations. An agent is cast as an entity which partitions possible futures by its choices, a rational cognitive agent and a rational reactive agent may make similar choices and may be viewed as being of the same agent equivalence class despite having different operating methods. Beliefs and knowledge are in the coaching domain and the strength of evidence for a particular action bringing about A may be reflected in the strength of a coaching agent's beliefs about that action.

Definition 14 *An agent* believes *a proposition if, as a result of its action or actions, that proposition may be true in some future world.*

Definition 15 *An agent* knows *a proposition if, as a result of its action or actions, that proposition will be true in all future worlds.*

This idea that knowledge is a true belief is widely used (see, for example, [41]) and can be easily represented by the *knowledge axiom* (see section 2.6.6). The importance of Fagin's observation, mentioned in section 2.6.1, now becomes apparent, the accessibility relation embodies an agents knowledge, beliefs and abilities.

2.6.4 More than one agent.

Consider two agents, α and β meeting at some location in their world. Figure 2.3 illustrates their potential interaction as a possible worlds structure. This illustrates some interesting aspects of multi agent interactions, it is clear to the observer who has an external vantage point (and from the way that the diagram represents the situation) that worlds W_2 and W_3 are equivalent. This equivalence may not be apparent to the agents

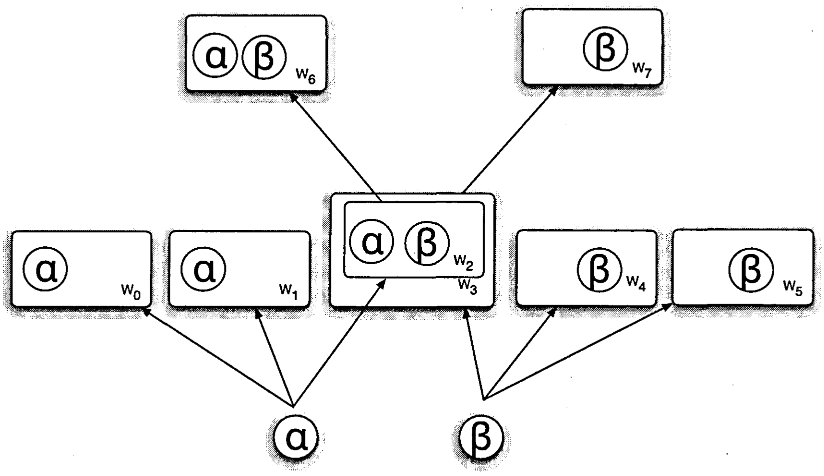


Figure 2.3: Interaction between two agents represented as possible worlds

because their view of the world is from within and they may use different symbols to represent the same things. The differences in the agents *beliefs* about the world is also apparent, α believes that W_6, following W_2 is possible. β considers only W_7 following its W_2 equivalent, W_3. An external observer has a privileged position and may see equivalences that are not apparent to situated agents.

2.6.5 Truth and validity.

Possible worlds semantics provide a framework for describing modal truth and validity from both a general and an agent centred perspective. This is a powerful property as it allows us to take an agent agnostic view, one which treats agents as autonomous choice partitioning mechanisms, of behaviour without weakening any of the aspects of the agenthood of the agency of actors within a society. Consider the valuation of the statement $\phi \land \psi$ at some world w, $V_w(\phi \land \psi)$, this statement is *true* just when $V_w(\phi)$ is true and $V_w(\psi)$ is also true. The worlds play no part in the valuation of the statement since it is truth functional and dependent solely on the values of ϕ and ψ. Consider the same propositional statement with a modal qualifier, $\diamond(\phi \land \psi)$. How can we say if this is *true* or *false*? For this to be true there must be at least one world in the set of possible worlds where $V_w(\phi)$ is true and $V_w(\psi)$ is also true. Similarly, for something to be necessary it must be true in all possible worlds. This gives evaluation rules for necessity and possibility:

$$V_w(\Diamond\phi) = \top \text{ iff } \exists w' \in W \text{ such that } wRw', V_{w'}(\phi) = \top$$

$$V_w(\Box\phi) = \top \text{ iff } \forall w' \in W \text{ such that } wRw', V_{w'}(\phi) = \top$$

The truth of a modal sentence is not dependent on the present value of any propositions that it contains.

2.6.6 More formally, some set theory and axioms.

Now that we have an informal notion of modal logic, an outline of Kripke structures and their associated sets of worlds and an outline of validity we can examine things in a little more detail. This section introduces the set theoretic properties of the relation element contained in a Kripke structure the simplest system, K and describes the properties of *normal* and *non normal* systems. Set theoretic properties are described before introducing a number of well known modal systems. The section finishes by informally outlining how these elements relate to agent systems.

Recall that the binary relation, R, of a Kripke structure controls which worlds are accessible from a given world. R may exhibit set theoretic properties [30, page 80] characterising the relation as being:

- Serial or extendible iff for every w in W there is a w' in M such that wRw'
- Reflexive iff for every w in W, wRw
- Symmetric iff for every w and w' in W, if wRw' then $w'Rw$
- Transitive iff for every w, w' and w'' in W, if wRw' and $w'Rw''$ then wRw''
- Euclidean iff or every w, w' and w'' in W, if wRw' and wRw'' then $w'Rw''$

Modal logic systems are generally accepted as having at least two basic properties [30, page 6], the distribution axiom, sometimes known as K, and the rule of necessitation.

These combined with propositional logic result in the most simple normal modal logic system, known as K (for Saul Kripke). Additionally all *normal* systems of modal logic contain the modal validity Df$\Diamond$[4] and the rule of inference RK[5] [30, page 114]. A normal modal logic system is simply an extension of *K* and is usually obtained by adding axioms or properties which *strengthen* the logic. The truth of a proposition interpreted in a K system is preserved in other normal systems. Non normal logics are weaker than *K*.

These properties together with necessity, possibility and propositional logic allow the derivation of a number of axioms and the S5 system of modal logic[6]. S5 is generally regarded as a logic of knowledge [40] and its axioms are listed below with brief descriptions of how they relate to an agent's knowledge:

- K. $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$ If a conditional and its antecedent are both necessary then so is its consequent, this describes closure under consequence.

- T. $\Box\phi \rightarrow \phi$: Whatever is necessary is so. Often referred to as the knowledge axiom, it treats knowledge as true beliefs and provides an agent with a means to distinguish between what it knows and what it believes. T is characteristic of reflexive relation functions.

- 5. $\Diamond\phi \rightarrow \Box\Diamond\phi$: What is possible is necessarily possible. The *negative introspection axiom,* an agent knows what it does not know. 5 is characteristic of euclidean relation functions.

- D. $\Box\phi \rightarrow \Diamond\phi$: Whatever is necessary is possible. This ensures that agents do not hold contradictory beliefs or knowledge, an agent cannot believe that today is Tuesday and simultaneously hold the belief that today is Monday. D is characteristic of serial relation functions.

- 4. $\Box\phi \rightarrow \Box\Box\phi$: whatever is necessarily so is necessarily necessarily so. The *positive introspection axiom*, an agent knows what it knows. 4 is characteristic of transitive relation functions.

[4]Df$\Diamond$. $\Diamond\phi \leftrightarrow \neg\Box\neg\phi$ what is possible is just what is not necessarily not so. Possibility in terms of necessity.

[5]RK. $\frac{(\phi_1 \wedge \ldots \wedge \phi_n) \rightarrow \phi}{(\Box\phi_1 \wedge \ldots \wedge \Box\phi_n) \rightarrow \Box\phi}$ $(n \geqslant 0)$ expresses general modal consequence, a proposition is necessary if it is the consequence of a collection of propositions each of which is necessary. For further details see [30, page19].

[6]C.I. Lewis proposed five systems of modal logic which he named S1 to S5, see Lewis and Langford, *Symbolic Logic*, 1931.

These are, but for one exception[7], independent and may be combined [93, page 39] as required. By adding these properties to the basic modal system K a family of logics can be obtained. Historically, $K_{Reflexive}$ is known as T, K_{Serial} is known as D, $K_{Reflexive+Symmetrical}$ is known as B.

As an agent learns (or is told) more about its environment and the results of its actions then it can tune the appropriate parts of its relation function so as to improve its beliefs of the accessibility of future worlds.

The standard $S5$ system presents a number of difficulties, difficulties that are exacerbated by the bounded nature of agents. Axioms 4 and 5 imply that the agent has perfect knowledge about what it does and does not know, there is some concern about the suitability of these axioms for a resource bounded agent [121, page 276] but it is accepted that positive introspection is less demanding and, perhaps, more suitable for such agents. Chellas [30] notes that, in addition to the standard axioms, $S5$ assumes two rules of inference. These are Modus Ponens and the rule of necessitation, *RN*, which means that $\vdash \Box A$ whenever $\vdash A$, Chellas [30, page 14] writes this as: $\frac{A}{\Box A}$. Wooldridge [121, page 275] notes that *RN* implies an agent's knowing all valid formulae, this necessitates the agent knowing an infinite number of propositional tautologies, a feat which is not going to be easy even for an unbounded agent. Additionally, K implies that an agent's knowledge is closed under consequence, this taken together with RN give rise to the *logical omniscience problem.* K and RN mean that an agent cannot hold logically inconsistent beliefs. Closure under consequence of a set of inconsistent formulae is the set of all formulae so an agent holding inconsistent beliefs must believe all formulae. Since K and RN are at the heart of any modal logic system then so must the logical omniscience problem, this is not good for bounded agents. Konolige (see [121, page 276]) suggests that logical consistency is much too strong a requirement for resource bounded reasoners and suggests non contradictory reasoning as an alternative. An agent cannot simultaneously believe ϕ and $\neg\phi$ but could hold logically inconsistent beliefs. Such a weakening may not cause any problems for an agent, just as an agent is bounded the results of its actions may well have bounds.

With allowances for logical omniscience a modal logic could feasibly be used to form the core of a resource bounded agent's reasoning system. Approaches to dealing with logical omniscience include

[7] If a relation is reflexive then it is also serial, consider $\phi R \phi$, ϕ certainly relates to something.

Levesque's logic of implicit and explicit belief (see [122]). Levesque proposes an agent with a small set of explicit beliefs, a larger set of implicit beliefs and logical operators for each. Explicit beliefs are represented semantically by a weakened possible worlds model and implicit beliefs by a standard possible worlds model. This approach brings other problems, most notably an inability to represent nested beliefs which may render it incapable of dealing with common knowledge (see [54] for a description of common knowledge) which may be a triggering force in agents co-operating to investigate their environment.

2.6.7 Why use modal logic?

There are alternatives to modal logics for representing agent systems. Petri nets, for example, were devised as a tool for modelling a particular class of problems [90], the class of discrete event systems with concurrent or parallel events. This encompasses the class of problems commonly encountered in distributed systems. Agent systems have much in common with distributed systems so tools suited to modelling distributed systems ought to be at least useful in modelling agent systems.

Non deterministic environments are easily represented by Petri nets, if more than one transition is enabled then the choice of which transition to fire can be made in a non deterministic manner. This choice could be random, controlled by forces that are not modelled or controlled by an agent. This non determinism is useful in modelling agent systems but can introduce significant complexity. The generally accepted way of dealing with this is to consider the firing of a transition to be an instantaneous event. Since time is continuous and instantaneous events require zero time then the possibility of two events occurring simultaneously is zero. Tasks that have a time requirement can be decomposed into discrete, instantaneous *start task* and *finish task* events.

Holvoet [65] concludes that using a single Petri net for modelling a large and complex system is not really feasible. Such systems would require the use of a number of dissimilar nets which may lead to unnecessary complication. Social systems and complex heterogeneous systems appear to be unsuitable candidates for Petri net representation.

Halpern [52] notes that logic has permeated computer science over the last thirty years and that computer science has benefitted from an extensive and continuous interaction with logic. Modal Logic has developed into a powerful mathematical discipline dealing with restricted description languages for talking about relational structures [36]. In the *Advances in Modal Logic Initiative* start-up document de Rijke [36] notes that in addition to considerable theoretical advances there has been a rapid expansion of the use of modal logic in computer science, cognitive science, linguistics and philosophy.

Modal logic and "classical" logic are systematically related [10] so that by using a set of standard translation rules it is possible to translate a modal sentence into a first order logic sentence. Some of these standard translation clauses are (after [10]):

1. $ST_x(p) = P(x), p \in PROP$[8]

2. $ST_x(\phi \to \psi) = ST_x(\phi) \to ST_x(\psi)$

3. $ST_x(\Box\phi) = \forall y(Rxy \wedge ST_y(\phi)$

The modal sentence $\Box p \to p$ can be translated by way of these clauses:

$ST_x(\Box p \to p) = ST_x(\Box p) \to ST_x(p)$	Clause 2.
$=ST_x(\Box p) \to P_x$	Clause 1.
$= \forall y(Rxy \to ST_y(p)) \to Px$	Clause 3.
$= \forall y(Rxy \to Py) \to Px$	Clause 1.

If something can be represented in modal logic then it can also be represented in first order logic. So why use a modal logic? Blackburn [10] notes that although modal logic and classical logic talk about the same models they do so in different ways each providing different *meta-logical* properties. Modal logic may provide both a global and internal perspectives on models which, as we noted in section 2.6.5, is a powerful tool for dealing with agent systems. Recall that agents are bounded and may only have partial access to

[8] *PROP* is a given set of propositional symbols, $\{p, q, p\prime, q\prime, \ldots\}$

the environment (section 2.2.2) and that an agent system may contain a number of agents (definition 1). Halpern [51, page 196] notes that multi agent systems may be represented by considering each agent's *local* or *internal* state which encapsulates all if the relevant information that that agent has access to. Modelling a multi agent system requires modelling at an agent level and at a system level. Modal approaches treat local and agent internal perspectives as just that – the perspective of a single, situated agent. The reasoning framework of modal approaches is the same at this local level as it is at a system level, such approaches allow for a reasoning framework that is consistent across the spectrum from single agents to agent systems.

2.7 Logics of agency, the STIT approach

We have seen that modal logic is a useful and powerful tool for representing the evolution of a world from both an observers perspective and from a situated agent perspective. We have seen that this allows us to view agents in an agnostic manner treating them as autonomous mechanisms which partition the future according to their choices. We now link this view more closely to the agents and agency aspect of the notional partitioning mechanism.

One of the main themes of logic of agency is that actions are abstract and that they may be identified with what they cause. The first semantics for this were laid out by Chellas [29]. Two main groups of logics of agency then followed, Kanger [69] and Pörn [91] [92] adopted a *bringing it about* approach with Belnap and Perloff [34] and Horty and Belnap [67] adopting a *seeing to it* approach.

The bringing it about approach was occasionally adopted but there were problems with the clarity of the semantics.

2.7.1 Agency and ability in a modal context

The earlier treatment of agency dealt with the *possibility* that an agent may bring something about. This views agency in an impersonal and abstract manner which is more suited to a systems than a society centred view. It captures the fact that an agent may be involved in bringing something about but does it genuinely represent how an agent's *ability* to bring something about led to it being brought about. Possibility is a vital aspect in any treatment of agency. However this must be bounded so as to ensure that limited agent resources are not "wasted" on frivolous exploration of possibility. Belnap notes [7] this by describing *objective possibility* and Xu [124], perhaps more concisely talks of *possibilities based in reality*. Xu notes that we use the terms *possibility* and *possible course of history* in a manner that does not encompass everything that we can conceive of, factual possibility rather than fictional possibility. Belnap lists a number of points [7] indicating the need for factual possibilities in a treatment of agency;

- There is no probability without possibility.
- There is no action, no doing, no responsibility.
- There is no agency without possibility.
- There is no causality without possibility.

Possibility appears to be an integral part of agency and any agent wishing to learn about its environment must be able to identify its role in possible events.

Belnap outlines several principles of agency [7], if agency is to be attributed to an entity, α, then, in an appropriate context, we must be able to find a sentence saying that α *sees to it* that something is brought about.

There is no agency without choice, our aim is to develop agents and not simple, situated automata that react only in a stimulus and response manner. This has a number of implications, if an agent is tasked with ensuring that A does not hold and it has an ability to counter A and that ability is rechargeable and transferable. When A occurs that agent may face the choice of running to a charge source repeatedly in an attempt to counter A or forming part a bucket brigade like chain to transport the counter A ability. This choice of methods endows it with agency. An important aspect is that an agent cannot make tomorrow's choices today, an agent cannot position itself somewhere and wait for A to occur as this prevents it from doing other things. Any treatment of agency and ability should be able to accommodate these principles.

2.7.2 stit theory

The idea of representing an individual's agency or actions as a modality is an old one, one which dates at least as far back as 1100 ad and Anselm of Canterbury [67]. The notion of agency as a modality idea has also been dealt with by Anderson, Åqvist, Chellas, Fitch, Horty, Kanger, Pörn, Segerberg, von Kutschera and von Wright. This idea has received attention more recently with a series of papers by Belnap and Perloff using an approach which they termed *stit theory* and which captures the intuitive concept of an agent being able to *see to it* that some state of affairs is brought about. Stit theory expresses this notion by using a construct of the form $[\alpha\ stit\colon A]$ which says that agent α can *see to it that* that A is brought about. We use this concept of "seeing to it" to characterise the abilities of an agent and in a bounded environment and coaching agents will operate on this notion of agent ability when reasoning about agent behaviour.

Belnap and Perloff [34] coined the term *stit theory* to describe their approach to treating agency as a modality. Stit theory is predicated on the notion of agent ability and ability is characterised as the agent being able to see to it that either something is done or a state brought about. This is usually expressed in an abbreviated form as:

$$[\alpha\ stit\colon A] \tag{2.6}$$

This notation represents that notion that agent α has the ability to bring about A. Let us consider Kenny's example of a poor darts player [72] used by Horty and others.

In figure 2.4 a darts playing agent, α, has the choice of throwing or not throwing a dart. α*'s*, set of choices is represented at that moment, m, by the notation $Choice^m_\alpha$ which is standard in the literature. If the agent elects to throw a dart then both possible outcomes guarantee hitting the board but the agent is not sufficiently skilled to be able to guarantee hitting either the top half or the bottom half of the board. If the agent elects not to throw a dart then the only possible outcome guarantees that it does not hit the board. It is obvious, in this case, that the agent has the ability to see to it that it hits the board but it cannot guarantee hitting the top or bottom half of the board though either is clearly possible. Similarly, we can say that the

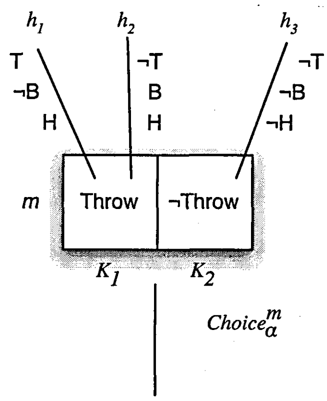


Figure 2.4: The choices available to a poor darts player

agent knows that by throwing a dart it can hit the board, the agent can only believe that it can hit the top half of the board. If we take H as being "hits the board", T as "hits the top half of the board" and B as "hits the bottom half of the the board" then it is easy to see that [α *stit*: *H*] is true, the agent has the ability to *guarantee* that it can hit the board. Similarly, [α *stit*: *T*] and [α *stit*: *B*] are false. The agent does not have the ability to guarantee that it will hit the top half of the board

An attractive aspect of this analysis is that, in broader contexts, it is not prescriptive, if it is cold then an agent may have the ability to make itself feel warmer by either putting on a sweater or turning on the heating. This method agnosticism is especially attractive in agent systems where agents with differing abilities may be able to bring about the same results. We can state goals for agents and whilst we are concerned with results we can remain disinterested as far as methods are concerned. The use of such an analysis for normative systems is especially attractive.

Horty [67] notes that Belnap and Perloff give an intuitively compelling account of their *stit* operator within a logical framework of indeterminism, a framework which holds possible future worlds that an agent may or may not have experienced.

2.8 Agency and what it means to *see to it that...*

Our system is concerned with creating reactive agents that operate in a social manner so as to satisfy certain norms imposed by a societal rules and possibly observed by an external entity. These norms are characterised as safety and liveness properties (see section 2.4.1) and we consider them as being "strict" norms. Reactive agents, as their title implies, don't really deliberate their actions. At its most simple their operation is a cycle of sensing their environment, selecting a suitable action from a set of available actions and then carrying that action out. The agent's main concern is selecting the appropriate action for its current percepts. This sense–select–act cycle has a very short horizon but this does not mean that a reactive agent is incapable of planning so as to bring about specific goals and neither does it mean that reactive agents are incapable of learning or being coached.

Horty's analysis of agency deals with what we described earlier, in section 2.2.7, as a *strong notion of agency*. This is the type of agency which may be exhibited by very able agents with strong cognitive and planning abilities. Our agents are rather less able and a full application of Horty's analysis in unnecessary. The differences between Horty's analysis and our needs is most apparent when it comes to refraining from acting where our needs are very much simpler.

Consider a world – possibly bounded, this is of no real significance, with a population of agents. The world is governed by a simple norm, *A should not hold.* Unfortunately instances of A will spontaneously appear in this world so this norm obliges agents to counter A when they come across any instances of A. The DSTIT operator implies that an agent has the ability to deliberate about its actions – here is an instance of A and I am able to counter it. Do I counter it or do I wander off and do something else? Not countering A would be a violation of a system safety property so there is no real need for deliberation unless, of course, the agent would be damaged in the process. The existence of the norm, $\neg A$, may precede the agents ability to satisfy it.

In addition to actor agents the world contains a number of coaching agents. Coaching agents are a separate class of agents that can not manipulate the environment directly but can analyse observed histories of events left by actor agents. Coaching agents pool the observations of numerous agents and attempt to understand the abilities of these agents so that they can begin to suggest patterns of behaviour that may satisfy system norms. The coaching agent's analysis of observed events – the link between observed histories and improved behaviours – is the motivation for this work. We seek to provide a logical characterisation of such an analysis.

2.9 STIT and belief

Stit semantics may be used to characterise beliefs about the behaviour of agents. This is usually done in cases given strong positive introspection, environment knowledge, knowledge of abilities and the consequences of actions. We are more interested in simple reactive agents and, consequently, use stit semantics as a tool for analysing agent influences on the environment. This gives a tool which we may use to guide a coaching agent's *beliefs* about agent ability. Beliefs are not absolute and, consequently, we represent them by a possibility or "may lead to" operator. Moving to the observer's point of view, if an observer sees many cases of an agent choosing a particular action that leads to $\neg A$ and a few cases of counter evidence then the observer may hold that that action leads to $\neg A$. With weaker evidence the observer may hold a belief that the action *may lead to* $\neg A$.

Because this stit operator is based on belief of an agent's ability rather than the truth functional status of formulas in a possible world we need to adopt different evaluation criteria. The statement that an agent's action *may lead to* A is true if and only if there is some point where the agent has a future choice that brings about A.

2.9.1 Belief as a ternary operator?

When does a coaching agent believe that [α action may lead to: A]? In order to believe this then a coaching agent must have evidence that α is able to bring about A and must have no evidence that α can not bring about A. If α has managed to bring about A in the past and a recent attempt has failed then there is both positive evidence and counter evidence against its ability which leaves it unsure of its absolute ability in a given set of circumstances.

We have considered using a ternary settledness operator to accommodate this. Using a minimal model, in the same manner as Brown [25] addressed the problem, brings problems and would require a higher order modal logic to deal with potentially complex relation functions. A ternary operator is a little less "brutal" in partitioning future worlds than a normal true/false operator A ternary stit type operator may be evaluated as:

- Settled true if α has evidence that it can bring about A and has no evidence to the contrary.
- Settled undecided if α has evidence that it can bring about A and also has evidence to the contrary.
- Settled false if α has no evidence that it can bring about A.

This ternary evaluation may be thought of as expressing a coaching agent's confidence in an actor agent's abilities. Indeed, this ternary evaluation is a prototype for our coaching agent's hypotheses and this is something which we shall examine in greater detail in later chapters.

2.9.2 Belief and knowledge

This work is based on the notion of agents being able to *influence* their environment and that this influence is extendible across agents. The consideration of belief as a ternary operator provides one of the foundations for a theory of influence, the *belief* that an agent can bring about A. It may be that an agent can not help but to have A hold after an action. A may be a constant and in such cases it is clear that an agent has no influence. This provides a second foundation for a theory of influence, in order for us to be able to say that α has influence over A it must be that A is not a constant. Considering this following the definitions of belief and knowledge in the context of agents, definitions 14 and 15 respectively, allows the formation of conditions for agent influence. This neatly partitions into knowledge and belief, in order for us to be able to say that an agent has influence on a proposition we must have knowledge that that proposition is not constant and evidence that an action by the agent results in that proposition holding. The knowledge content is not that a proposition will change in all future worlds but that a proposition is changeable and to this end we need only observe one instance of change.

2.10 Other studies of influence

We noted earlier that other researchers have used the term "influence" in agent systems research. Ferber and Müller's multi agent based simulations (MABS) [44] formalism uses the notion of agent influence. Ferber's model is an action theory which is based on a foundation of influence and the environment's reaction to that influence. Agents do not *act* in the traditional sense, instead they generate influence which may or may not result in an effect on the environment. Instead, there is a distinction between influences and reactions in a system that is composed of two sets of dynamics. An agent may *attempt* to lift an object or *attempt* to counter an instance of A and if its attempt has the appropriate influence then it succeeds. This differs from our work in many respects, the most notable being a mechanism which aggregates agent influence at a global level and this is a level of operation that we are keen to avoid. Ferber's approach is implementation dependent and does not readily model simultaneous actions (see Michel [86], indeed the original theory handles joint actions in a way that casts a single agent as an initiator and has it making decisions for other agents. Michel [86] notes that the Ferber and Müller model has not been applied in its natural form. Weyns [119] notes that Ferber and Müller's approach is limited to synchronous systems. Michel [86] goes on to describe an approach which applies Ferber's influence / reaction model and goes some way to addressing these problems. Although the notion of influence is closer to Ferber's Michel does adopt a two step process consisting of an influence phase and a reaction phase. This agent action and environment response is something that we address implicitly in our underlying branching time model and instantaneous view of actions discussed in chapter four. Although we model our system on a discrete branching time framework this is as a convenience for modelling a logical system. The time steps in our system are of an arbitrary size and may be arbitrarily small. Our notion of influence is similar to that of Ferber and Müller's but we are working in an opposite direction taking the results of an action and attempting to divine the influences that caused it. Further, we are dealing with influence as an "enabling token" which may be used as a vehicle for analysing and synthesising agent behaviours and this is a very different approach from Ferber and Müller's simple aggregation of causes.

2.11 Computational tractability

We have informally characterised agents, in a branching time setting, as mechanisms that partition the future by their current choices. Agent choices are associated with *moments* in the branching time framework but not all moments provide choices – some may be vacuous where an agent has no choice that influences its future. For long term goals, agent level or system level, it is preferable to have an enumerable set of choice points. Having a non enumerable set of choice points may force the agent into a sequence of *constant decisions* which is not good. Belnap et al. describe a "busy chooser" as an agent that makes infinitely many choices in a finite period. This leads to scenarios where there is no single witness. A witness, in this context, is a characteristic of achievements STIT expressions and is a moment where a prior choice ensures that things evolve in such a way that A is guaranteed at the evaluation instant (Belnap et al. [8, page 36]). Instead, witnessing is by a chain of moments (Belnap et al. [8, page 249]). The difficulty with chains is that there may be no last member of the chain. Xu [123] proves a correspondence between the numbers of modes of actions/inactions and the complexity degrees of busy choice sequences and this avenue may provide an insight into the optimisation of sequences of behaviours.

The *ten minute mile* example used by Belnap et al. [8, page 49] illustrates the busy chooser difficulty with STIT semantics. This is illustrated in figure 2.5, at m_0 an agent sets of to run a ten minute mile. It faces a choice, run at a pace that allow the agent to complete a ten minute mile, A, or at a pace that will not, $\neg A$. This is not a one time decision as it is immediately followed by a similar choice which is immediately followed by a similar choice and so on until the agent completes the mile or runs beyond ten minutes. At any point in the run the agent faces the choice of continuing to run at ten minute mile pace or not. This makes the runner a *busy chooser*, there are an indeterminate number of decisions between m_0 and the agent completing its ten minute mile.

This presents obvious difficulties for computational solutions. Something more "granular", something with enumerable and discrete decision points would be much better. We see the $[\alpha\ stit\colon [\beta\ stit\colon A]]$ construct as having an enabling moment. In this example it's where agent α makes some choice which *enables* β so that it may see to it that A holds.

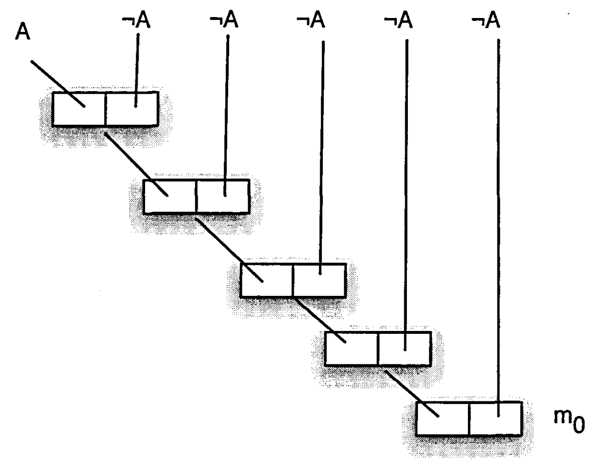


Figure 2.5: The busy chooser, a chain of choices for a ten minute mile

2.12 Self organising systems and emergent behaviour

Our research has elements of self organisation and emergent behaviour. Intuitively the system is expected to be self organising as it is intended to adapt its behaviour and –save for an external observer which evaluates target achievement – there are no external agencies involved in guiding or evaluating system progress. Self organisation in agent systems appears to contradict the second law of thermodynamics – thermodynamic systems are often used as examples of self organisation. Van Dyke Parunak and Brueckner [117] note that this has been explained by couplings between the macro and micro levels of a system. Our approach explicitly admits this view by using coaching agents as the macro component and actor agents as the micro component with a random element in actor behaviour providing entropy.

This work builds on earlier experiments with pheromone driven emergent behaviour where Logie et al. [82] investigated normative descriptions of pheromone based reactive systems (Van Dyke Parunak and Brueckner [117] note that the pheromone in such a system is the micro level with agents operating at a macro level). We make no claims for new methods of identifying and predicting emergent behaviour, our approach is to explore a behavioural state space by concentrating on areas where influence is evident and ignoring areas of the state space where agents have no influence.

2.12.1 Self organisation

The notion of self organisation occurs in many fields, Gershenson [48] lists cybernetics, thermodynamics, mathematics, information theory and synergetics. Self organising systems also occur in nature, animal flocking and herding for example. Humans provide many examples self organising systems, people moving through a crowded area or in and out of a crowded commuter train for example. As with agent systems there appears to be no single definition of what a self organising system is but Gershenson and Heylighen [48] note that one property is often used to characterise self organisation, that of *negative entropy*. Since we intend to use notions of influence in our system we may use this as a metric for entropy, our system becomes more ordered as agents use their influence less wastefully. Heylighen [61] indicates that another definition for self organisation is that it is the creation of a globally coherent pattern out of local interactions. Global coherence, in the context of our system, would be agents not violating system level norms which are defined in a procedure agnostic way. The self organisation is implicit in that we expect the system to learn on its own with no intervention from external agencies. Coaching agents will assist actor agents but these are internal elements of the system and, beyond knowing system level norms, are free from external influence.

Heinz von Foerster developed the notion that noise within a complex system is a driver for organisation, this counter intuitive notion was dubbed the "order-from-noise principle" (see, for example, Heylighen [61] or Foerster [118]). This notion is predicated on "attractors" and the idea that the larger the random perturbations, noise, that affect a system, the more quickly it will self-organize or produce *order*. Our approach builds on this notion, coaching agents observe influence and then reinforce it in a manner similar to positive feedback by seeding influence maximising behaviours into the environment.

2.12.2 Emergent behaviour

Heylighen [61] notes that organisational closure may turn a collection of interaction components. Heylighen [61] notes that many natural systems exhibit *emergent behaviour*:

> Organizational closure turns a collection of interacting elements into an individual, coherent whole. This whole has properties that arise out of its organization, and that cannot be reduced to the properties of its elements. Such properties are called emergent.

This may be thought of as a subset of self organising behaviour. A flock of birds may fly in such a way that there are no collisions, this is simply a scaling of individual behaviour. A single bird may be able to carry out the same journey without the support of a flock. Emergent behaviour is, as per the quote from Heylighen above, best thought of as aggregate behaviours that individual entities are incapable of. A termite nest, for example, could not be built by an individual termite and food foraging may be carried out by single termites or groups of any size.

2.12.3 System or society

Dastani et al. [35] offer an interesting perspective considering individual agents and multi agent systems as different computational entities. Their approach is based on co-ordination operations and has a well defined interaction protocol. We adopt a different approach without explicit co-ordination and this allows us to build systems with simple, reactive agents.

Jennings [68] notes that a major selling point of purely reactive agent systems is that overall behaviour emerges from interactions between component behaviours. We are more concerned with *societies of agents* than the interaction of behaviour components within agents. Our approach, in this work, is to deliberately keep agents simple so that overall behaviour emerges from interactions between agents rather than interactions between agent components.

Chapter 3

Agent influence

Our outline of STIT theory indicated that it was a useful tool for representing agency. Our approach is founded more in computer science than in general agency and, for us, this introduces constraints. These constraints are, for the most part, the introduction of bounds. Our agents are tightly bounded in their abilities (both physical and cognitive), perceptions and memory. The agent world is similarly bounded being a finite construct containing a bounded number of bounded entities. The import of these constraints is that the nature of computational agents and their environment makes experimentation feasible and this is the path which we intend to take.

Our work is intended to address difficulties presented by standard STIT theory, these are concerned with what are called *other agent nested* sentences, sentences which describe where a number of individual agents combine their abilities to jointly bring something about. Syntactically such sentences are well formed but present semantic difficulties. These difficulties centre on the nature of one agent *seeing to it* that another agent sees to it that something holds or is brought about and these difficulties centre on the *independence of agents* postulate (see Belnap [6]). Xu [125] summarises this intuitively and neatly:

> this constraint says that any combination of possible choices available to different agents at the same moment must be compatible. That is to say, roughly, at each moment, each agent can choose each of his alternatives, no matter what the other agents are doing at the moment. Thus

a possible choice for an agent at a moment should be considered as a real alternative for the agent, i.e., the realization of that alternative is exclusively up to the agent.

Xu's description of the independence of agents postulate, above, provides a convenient point for the introduction of the notion of agent choice in the context of this work.

Definition 16 *Given an agent, α with a set of abilities which are characterised as atomic actions, $Actions_{\alpha}$, we define α's choice set, $Choice_{\alpha}$, where $Choice_{\alpha} \in \mathcal{P}(Actions_{\alpha})$ and containing viable actions or combinations of actions such that all actions associated with a choice will be executed simultaneously.*

By "viable" in the above definition we mean that, at agent level, all of the combinations of choices are sensible in the context of the agent's world. For example, a choice element containing "move forwards" and "move backwards" is not viable for a mobile agent. As a convenience we extend the $Choice_{\alpha}$ notation of definition 16 by considering it at a particular moment.

Definition 17 *Given an agent, α with a choice set $Choice_{\alpha}$, and a moment, m, we write the the set of viable choices at that moment as $Choice_{\alpha}^{m}$ where $Choice_{\alpha}^{m} \in Choice_{\alpha}$ and contains viable choices at moment m.*

Viable, in the context of definition 17, encompasses the viable of definition 16 with the additional constraint that the set of choices are only those viable at moment *m* and be ordered in such a way as to group certain of them for convenience.

We have noted, in chapter two, that agency has a large element of autonomy and this is explicitly expressed by the independence of agents postulate. A strict reading of STIT seems to imply that one agent has control over another, if this is the case then the other is no longer an agent, this is something of a difficulty. In another of the touchstone phrases from the introductory sections Chellas [31] (and in Belnap et al. [8, page 275]) notes that it would be:

"...bizarre to deny that an agent should be able to see to it that another agent sees to something"

Xu [125] introduces the notion of *nested choice formulas* to overcome the difficulties arising from the independence of agents postulate. Xu, however, deals with current choices only and is consequently limited to the consideration of deliberative STIT constructs.Having frequently raised the notion of a "strict" reading of STIT at many points we now state a definition for this term.

Definition 18 *Given an agent, α, and a set of observations of α choosing a particular action, K, in a given set of circumstances identified solely on the basis of α's percepts. We say that α choosing K, α/K, has the unambiguous ability to see to it that that some proposition, A, holds if and only if the set of observations contains no cases of α/K leading to $\neg A$. We consider this unambiguous ability as satisfying a strict reading of* STIT. *The presence of observations of α/K leading to $\neg A$ fails to satisfy a strict* STIT *reading but does not necessarily mean that α has no influence over A.*

Applying definition 18 to multiple agents and nested agentives a strict interpretation of $[\alpha\ stit\colon [\beta\ stit\colon A]]$ requires that α has some means of guaranteeing that β sees to it that A holds. This is certainly possible but in cases where one agent is able to see to it that another agent sees to something there is usually a social or societal framework which governs agents by, for example, obligation or sanction. Belnap et al. approach this problem by considering various interpretations of other-agent nested STIT constructs. These are deontic, disjunctive, probabilistic and strategic. These interpretations bring additional system requirements which make them unsuitable for simple, reactive agents and draw us away from our aim to present a simple reading for nested other agent statements. A deontic interpretation, for example, requires mechanisms for creating and transferring agent obligations and, by extension, agents must be able to reason about obligations in order to operate in such a system. Although we are interested in such normative approaches this is unsuitable because our very simple agents have no means of either imposing obligations or assuming obligations from others. Probabilistic approaches are similarly unsuitable because we are more concerned with qualitative than quantitative aspects of agent behaviour. The disjunctive approach fails in a similar fashion to the deontic reading because simple agents cannot force choices on others.

Our agents operate in a society and we intend them to adapt their behaviour so as to build a society that conforms to a number of norms. These norms are cast as safety and liveness properties but they are described in an agent agnostic manner at "system level". Our approach to this is to maximise what we term *agent influence*, the observed and apparent ability of agents to singly or jointly bring about change in their environment. Our interpretation of nested agentives is similar to the strategic reading in that the strict STIT element is replaced by an evaluation based on the *influence* which agents may have on other agents and may be interpreted as a reactive plan or strategic interpretation of STIT. There are other approaches to managing strategies, van der Hoek et al. [116] describe a counterfactual extension to alternating time temporal logic which allows reasoning about strategies. This involves an element of agent commitment and is, therefore, more suitable for cognitive agents than our reactive actors.

Although this work has similarities to that of Belnap et al. [8, page271] there are a number of significant differences. The similarities are that we minimise the role of agent intention, and, consequently, the role of the individual agent, to concentrate on causality which we view in a social context. One difference in our approach is that we avoid a reification of actions, instead we are considering agent influence and we characterise nesting explicitly via agent actions. Our influence based approach implicitly admits witness by chains, as outlined in section 2.11, although we do so in a simple manner and this is described in section 3.5.6.

Our work is intended to be a component of a computational system and this brings both benefits and difficulties. The benefits are that we are able to generate experiments to test theories and, if necessary, to trace and characterise every aspect of an agent's behaviour. Consequently, the behaviour of any system composed of computational agents may be fully characterised. The disadvantage is that we must define both our agents and the world that makes their environment. The computational systems aspect of our work means that parts of the usual branching time / agent choice which underpins STIT theories need to be defined explicitly with a computational system in mind.

3.1 What is influence?

So far we have mentioned the notion of *influence* frequently, have mentioned it relative to STIT and have stated that it is not the same as the influence described by Ferber et al. [44]. What, then, is influence in the context of this work?

The foundation of our notion of influence is that an agent may act so as to bring about a change in its environment and that the results of that action may be contingent. The key point is that an agent may have an ability only in certain circumstances and where these circumstances are dependent on other agents standard

STIT theory does not extend across agents without the support of additional frameworks. Rather than state a definition of our notion of influence here we observe that:

Observation 6 *We observe that an agent, α exhibits influence over A if and only if A is changeable and that in a given set of circumstances α may acted so as it is capable of changing its environment to bring about A. This set of circumstances may be contingent on the environment state, α's state or actions by one or more so called* other-agents.

This gives an informal foundation from which we may build a theory of influence.

3.2 Why investigate influence?

In section 2.12 we stated that we make no claims for new methods of identifying and predicting emergent behaviour. Our approach, we indicated, is to explore a state which represents an environment and at least one agent. Here we outline informally how we hold that influence operates by considering a simple state space which agents may move through by executing actions. The entire state space is not necessarily accessible to a single agent, one class of agents may have the ability required to reach a given state whereas another class may not.

We view this as a privileged observer, we may see where good states are achieved and we may also see the trails of actions left by agents. Agents situated in the environment simply act and, as a result, move through the state space.

As we observe the state space we are looking for two things; agents reaching a good state (which is system level requirement that observers are aware of) and an agent behaviour or sequence of behaviours that allow agents to bring this good state about.

This carries prerequisites, the state space must contain at least one instance of the good state and that good state must be reachable either as a result of individual agent action or as the result of the actions of a set of agents. We cast the system as a normative system and task the agents with bringing about good states so the latter requirement is simply a statement of the *ought implies can* deontic identity. There are many approaches to state space searching which guarantee complete coverage and, by a given metric, finding the optimal solution or good state. State spaces may be large, meaning that complete searches may consume formidable resources. Heuristics may be applied to constrain searches but these also present difficulties, looking ahead in an unknown environment is difficult and makes evaluating heuristics problematic.

A randomly behaving agent may eventually reach a good state but it is unlikely in the extreme to reach the good state via an optimal path. If two or more agents need to interact in order to reach a good state

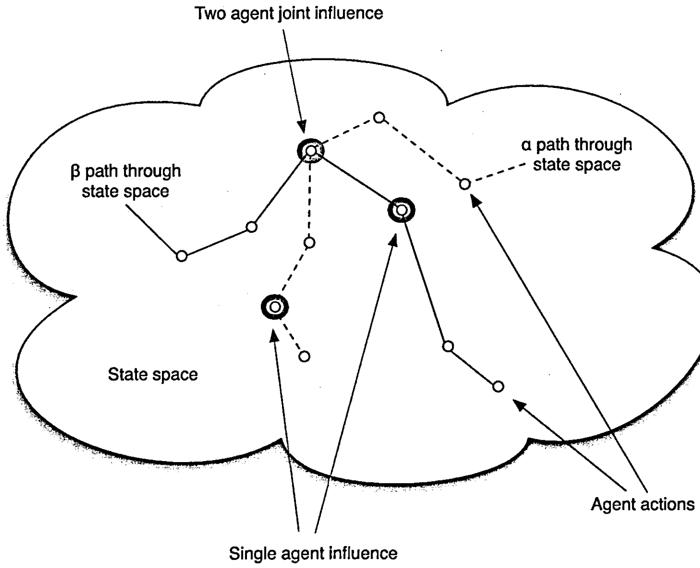


Figure 3.1: Two agent interaction in random state space walk

then the multiplicative combination of probabilities may greatly reduce the chance of an optimal solution. Figure 3.1 combines the situated agent and observer perspectives to illustrate this in a notional state space. Influential and non influential agent actions are indicated in figure 3.1, even though an action is not influential from the point of view of an observer it may still involve a change in the state space. An agent may carry out an action that does not change the observed environment but does alter the state space from that agent's point of view. β may attempt to lift a heavy object at state β:1 and fail. At state β:2 an observer may see β and the unmoved object, no change from the observer's point of view but β has expended considerable energy in its attempt to lift the object and perceives that it has moved to a different state. The observer is unable to see the internal change in β and may only consider this by inference from a series of observations rather than direct observation. However, the observed sequence of events provides an observed pattern that may be repeated in later observations. It may be that β can lift the object Milner [87] observes that:

> The behaviour of a system is exactly what is observable.

This is something which we shall consider in more detail in the following chapter but, for now, it serves to bring attention to the differences between observing a behaviour and knowing what drives that behaviour. An external observer may see some aspects of a system's behaviour but there may be internal aspects, such as β's reduction in energy, that are not evident. The coaching agent may be working with incomplete knowledge, Levesque and Lakemeyer [80, page 273] argue that a difficulty of dealing with incomplete knowledge is that it is computationally demanding. We wish to approach this in such a way as to minimise resource requirements so as to allow simple coaching agents to manage partial world knowledge.

There is an example of joint influence in figure 3.1, this may only occur where two agent paths cross (note that paths may only cross at points where agents act) in the state space. Paths may cross in such a way that agents are collocated and able to act simultaneously, α handing an object to β. There may be a temporal separation where one agent passes through the same location after the other agent and becomes involved in a sequential joint behaviour, α dropping an object for β to later pick up and use. From an agent point of view, β may be aware of the token that it has picked up but is not necessarily aware of the history of that token. The situated agent's view of its environment is limited to its local perceptions and the observer's view of the world is limited to what it can observe.

We introduce coaching agents as an intermediary between the observer and the actor agents. Coaching agents are privy to partial agent internal data and are able to aggregate this so as to construct a fuller representation of events in an environment. This aggregated local data means that coaching agents have a broader view of the environment than actor agents but they are still not privy to the global view that an external observer may have. The coaching agent's view of its surroundings – the environment and actor agents – is discussed more fully in section 6.6 where we explore details of coaching agent operation.

Returning, briefly, to the deontic identity, the observer has knowledge of system norms and implicitly requires that actor agents bring about so called good states. Since only actor agents are able to change the environment this means that they must be the drivers of change and that change is a result of their influence. If the environment is not in a good state then the only way that it can reach a good state is by change and this implicitly means as a result of agent action. We characterise this agent driven change as *agent influence* and hold that by maximising influence – promoting change in the environment – a system is more likely to reach a good state.

Coaching agents are not aware of "good" states in the same way that an observer is but they are constructed with two implicit operating principles.

- Agent influence is a driver of change in the system.
- Complex behaviours involving two or more agents increase agent coverage of the system state space.

One way of considering such joint behaviours is to view the state space as being partitioned into a number of behaviour domains. The simplest level is the set of states that single agents of all classes may have influence over. This is contained within the domain of states that require two agents to bring about and so on. As these domains extend outwards from the single agent domain they encompass increasingly complex behaviours requiring preparation or cooperation from multiple agents for influence to be expressed. States outside of the single agent influence core are properties of societies of agents. This notion of behaviour domains is illustrated in figure 3.2, intuitively behaviours which enable transitions between domains lie on the boundaries between domains. The right hand image in figure 3.2 illustrates possible agent choices at such a *gateway*. Here α may attempt to transfer a token to β, if β behaves in such a way that it accepts this token then its potential behaviours move it into the two agent behaviour domain. If β does not accept this token then it remains in the single agent action domain. Such gateway behaviour may be sequential or simultaneous.

We approach this as an iterative process and hold that, over time and as more of the system behaviour is observed, the ways that agents exercise their influence may be altered so as to bring their joint behaviour closer to optimal.

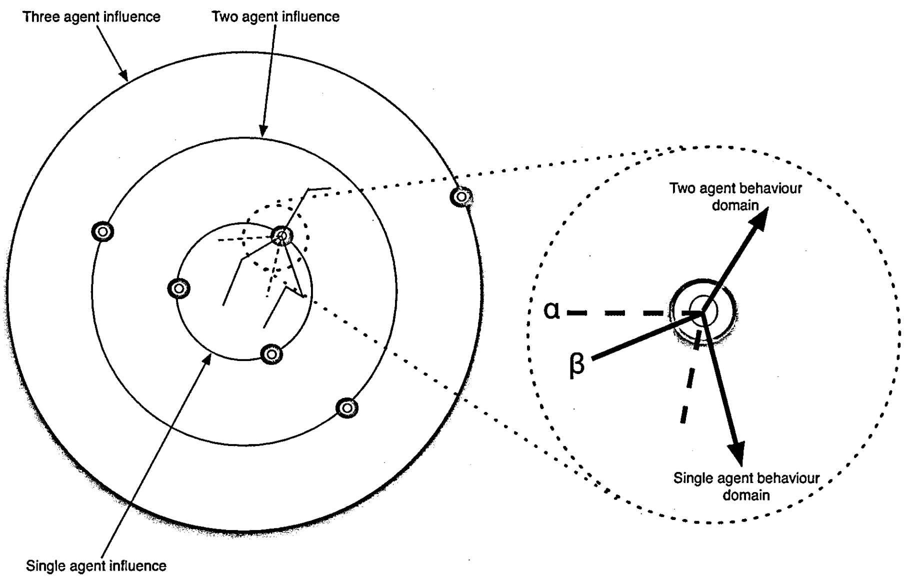


Figure 3.2: Gateways between single and multi agent influence domains

This thumbnail outline at an intuitive level indicates that influence gives us a means of managing agent ability in an uncertain environment and, more importantly, provides a means for coaching agents to explore areas of a state space that are only accessible to complex, aggregate behaviours. This drive towards complex aggregate behaviours will be a catalyst for emergent behaviours which allow systems to achieve states which an observer may view as goal states.

3.3 How to make sense of $[\alpha\ stit\colon [\beta\ stit\colon A]]$

Our notion of *influence* is based on agent ability, represented by *stit* notation. When we write $[\beta\ stit\colon A]$ we are stating that β has the ability to bring about A subject to the constraints imposed by the STIT operator embedded in the statement. At this point we consider a simple generic *stit* operator which assumes that an agent has a choice which will guarantee A at some evaluation instant. Our interest lies in cases where $[\beta\ stit\colon A]$ is contingent on a choice made by some other agent. For example, if α gives β an access token and β uses that token to enter a new behaviour domain by seeing to it that A holds when it was previously unable to do so. The fact that [β *stit*: A] holds is contingent on α having given β a token at some point before β executes the choice that leads to the truth of A. STIT notation allows for nesting so it would seem that this could be represented as [α *stit*: [β *stit*: A]]. Is this really the case for standard STIT operators?

When we say that [α *stit*: [β *stit*: A]] are we saying that α sees to it that β sees to it that A holds or are we saying that $\alpha's$ action makes it the case that β is *able* to see to it that A holds? If the former reading were true then α must exercise some influence over β but there is some concern that this is not correct. Belnap et al. [8, page 274], object to this reading on the grounds that it is inconsistent and justify this stance by demonstrating a contradiction. Assume that:

[α *stit*: [β *stit*: A]] is settled true at m_1, that w is a witness moment. (i)

Assume that m_2 is a counter such that:

[β *stit*: A] is not settled true at m_2. (ii)

By independence of agents we must have some moment m_3 such that both:

m_1 and m_3 are choice equivalent for α at w. (iii)

and,

m_3 and m_2 are choice equivalent for β at w. (iv)

By (i), (iii) and Chellas's witness identity lemma (definition 49), it must be the case that:

w is a witness for [α *stit*: [β *stit*: A]] at m_3. (v)

Taking (i) and (iii) with the positive condition we must have:

$[\beta\ stit\colon A]$ settled true at m_3 (let w_1 be a witness moment for this.) (vi)

From (v), (vi) and the witness identity lemma infer:

$w_1 \leqslant w.$ (vii)

This means that (iv) and (vii) imply (by backward monotony (definition 48) that:

m_3 and m_2 are choice equivalent for β at w_1. (vii)

The second witness lemma (definition 50) with (vi) and (viii) results in $[\beta\ stit\colon A]$ being settled true at m_2 which contradicts (ii) completing the proof.

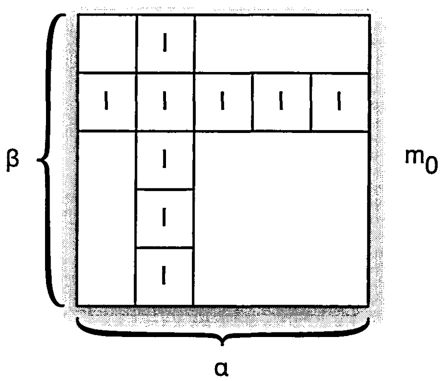


Figure 3.3: Illustration of the logical impossibility of $[\alpha\ stit\colon [\beta\ stit\colon A]]$ ($I = [\beta\ stit\colon A]$)

This is illustrated in figure 3.3 showing a two agent interaction. If we take I as meaning $[\beta\ stit\colon A]$ at witnessing moment m_0. Because I represents a *stit* by β wherever I appears in a choice row for β it must fill that row. This means that I is settled true at m_0 contradicting the negative requirement and settling the statement true at the witness moment. Alternative readings – deontic, disjunctive, probabilistic and strategic – are suggested. The notion of the deontic reading is that, in this set of circumstances, α creates an obligation on β to bring about A by the act of giving β some token or ability. The probabilistic reading indicates that α sees to it that there is a high probability that β brings about A. A disjunctive reading is offered but this requires that one agent has some power to force choices on another agent and is not suitable here. The strategic reading is perhaps most in keeping with what we are trying to do here, with a coaching agent in the system and working from agent ability represented by *stit* we are trying to develop a strategy for the system to use in order to bring about good states.

Returning to the token example of above, ideally if β receives a token from α it should hold on to that token until it encounters a set of circumstances where it may use the token to extend its influence by bringing about A. If β is predisposed to do this then α may *appear* to be seeing to it that β brings about A but if $\alpha's$ influence extends through β in this manner then it is at the cost of reduced choice for β. Our approach is to consider an alternative reading of this nested construct and that α *has influence* over β and that that influence plays a part in β seeing to it that A holds. Given that β is the agent which carries out that action that brings about A can we say that α *causes* β to do this? Xu [124] considers causation in branching time and describes conditions of necessity and sufficiency. The causing event is sufficient in *some* sense and necessary in *some* sense for the caused event. Xu is not concerned with what he terms *regularities*, that is one *object* followed by another and all objects similar to the first are followed by objects similar to the second. Instead, Xu concentrates on individual events and notes that we may speak of a particular event causing another without committing to notions of regularity. Xu's second reason is based on the claim that *The way particular things and events are related is the foundation of any regularity.* Regularity, Xu notes, is not a prescriptive rule that nature has to obey, but a descriptive pattern showing how particular things and events in nature are actually related. Thus even if the observers, standing outside of the world, try to figure out what the regularities in this world are, they must take into account the relations among particular things and events in order for them to succeed. We are attempting to find regularities by aggregating data on single instances of object sequences. Xu considers the chain of events leading to a house fire:

> Someone reconnected the electrical wires in a house in such a way that if a certain switch was later turned on, the house would catch fire. Another person, without knowing what the first person had done, turned the switch on. As a result the house caught fire. Assume that the particular turning-on of the switch was contingent before it happened. Many people may think, even under this assumption, that what the first person did caused the fire.

Xu intuits that it is the second person that *caused* the fire and not the first. What the second person did was sufficient for the fire under the circumstances, circumstances which include the first person tampering with the wires. Xu cites a second example to justify this intuition.

> Suppose that what the first person did was to hang a magnetized needle by an electrical circuit in such a way that the needle would start moving if a certain switch was later turned on, and what the second person did was to turn the switch on (assuming that it was up to him to turn the switch on). As a result, the magnetized needle in the magnetic field started moving.

It seems more reasonable to consider, in the second example, that the second person caused the needle to move. There is no element of blame in the second example and the two examples seem to be very similar if there is no consideration of blameworthiness.

Xu's account of causation seems to admit elements of intention and although this is of no interest to us it is worthy of mention as it preempts elements of coaching agent operation. Rather than looking for intention, coaching agents will be looking for connections between agent behaviours.

3.4 Extending notions of influence

An agent may have the ability to change its environment but in some circumstances this ability will be contingent on the choices made by another agent or agents. In these cases more than one agent contributes an *influence* to a final change. By examining agent behaviour and, in particular, agent choices and the results of these choices, it is possible to identify and measure such influence. In the following sections we informally explore the concept of agent influence as part of a complex behaviour. This informal exploration builds to an intuitive theory of influence which we formalise in the following chapter.

3.4.1 Sequential influence

We earlier considered two agents, α and β in a setting where α was able to provide β with a token which extended β*'s* influence. We introduce two new types of agent, γ and δ which are similar to α and β but do not rely on a collocated action for token exchange. The γ type agent may leave tokens in the environment and the δ type agent may, if it encounters one, pick the token up and to extend its influence. Figure 3.4 illustrates both agent's perspectives of such a sequence. The γ agent's history is depicted by a solid line and the β agents by a dashed line. At I_0 the γ agent has a choice of dropping or not dropping a token. δ has no influence at this instant and this is indicated by its inability at I_0 to choose between h_3, h_4, h_5 and h_6. If γ drops a token it follows h_2 and at instant I_a the diagram indicates that the histories (and not necessarily the agents) work' together in some way that extends δ*'s* influence in this purely sequential behaviour pattern. There is no need for the agents to be collocated as in the example above. The completion instant is clearly obvious but the enabling instant is not so easily identified. Enabling requires that γ drops a token bringing the environment into a state where δ may take advantage. The agents may not be collocated but some means of coordinating their behaviour seems to be required.

3.4.2 Joint collocated influence

Figure 3.5 illustrates β*'s* perspective of the world. In this example we assume that a token giver agent, α, attempts to give β a token at I_0 and the results of β*'s* actions at m_1 are based on this assumption. If β elects to move or use the token immediately then β will arrive at I_1, where there is a an instance of A, without the ability to counter A (we neglect the move choice at I_1 for clarity). This brief sequence indicates the influence that both agents contribute to bringing about a scenario where β is able to see to it that an instance of A is countered. In line with our informal outline of STIT evaluation it is possible for $[\beta\ stit\colon \neg A]$ at m_2 but not at m_3. It appears that β alone is unable to see to it that $\neg A$ holds. Although its action, given a token from α, is

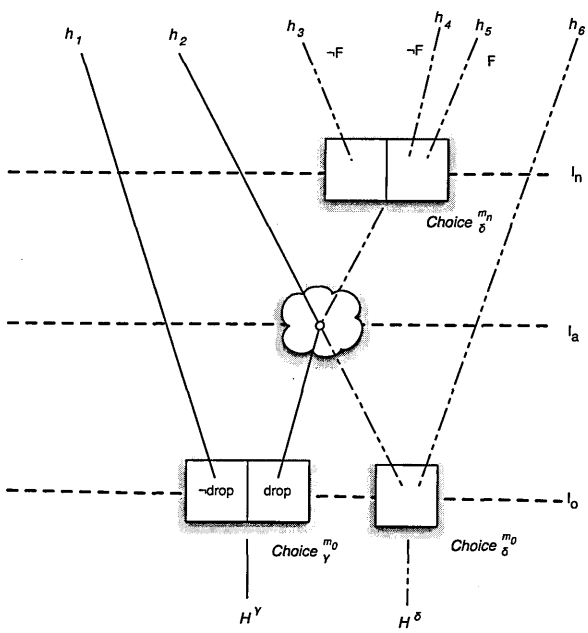


Figure 3.4: Sequential two agent action in branching time

necessary and sufficient to guarantee $\neg A$ It is also necessary that β be provided with a token beforehand and this requires the joint action with α at I_0 which we term an *enabling instant.*

Although figure 3.5 illustrates I_1 and I_2 as adjacent instants it is possible for β to carry out a number of actions between the enabling instant and bringing about A at the *completion instant.* The behaviour sequence requires that the agents be collocated at the same location in their world when the *influence extending event* – α choice which we describe as giving a token to β – occurs and a degree of inter agent cooperation may be required. β should not choose to do anything which inhibits α or the transfer of the token. Co-location is not the only condition under which influence can be extended.

3.4.3 Influence extending behaviour

These two simple examples reveal hints to some of the features of successful behaviour sequences and some of the difficulties of locating these behaviours. Each sequence has a starting point, in the example of section

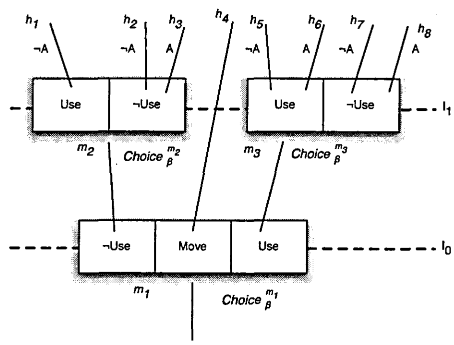


Figure 3.5: gent β's view of joint action where α gives β gives a token at m_1

3.4.2 this occurs at a meeting of agents but in the sequential example of section 3.4.1 the environment acts as an intermediary. The influence extending part of the behaviour involves bringing the environment into a state that potentially acts as a gateway to a new behaviour domain. This is evident the sequential example but it is also the case in the co-location example, from α's point of view β is part of its environment and α's attempt at giving β a token is an attempt to bring its environment into a state from which it is possible for β to move into a new behaviour domain. The starting point is an environmental transition to a *result possible* state but this may not be immediately evident.

The end point in each case is a readily identifiable transition to a good state as identified by system norms. Intuitively with less than perfect behaviours there will be more occurrences of starting points than of successful end points and many potentially good sequences will lead to unsuccessful results. There will also be occurrences of coincidental events over which the agent has no influence such as A spontaneously clearing itself coincidentally with a tokenless β attempting to counter A. Additionally, in a complex environment, good behaviours may be masked by a larger number of bad behaviours and this makes simple pattern searching impossible.

3.4.4 Uncertain histories

Here we note that notion of a history in this work differs from that of a history in the standard branching time framework. Semantically, the treatment of histories is the same in standard STIT and our theory of influence. Although semantic treatment is the same the concept of a history in our influence model differs in that it also admits uncertainty, this is necessary because of its accommodation of "other" agency. Were we to follow a strict STIT approach it would be necessary to consider sets of histories with the selection of which set is actually followed being dependent on choices by the other agent, allowing uncertainty means that we may operate on a single set of histories with the other agency being implicitly managed by the notion of influence.

3.4.5 Semantic extensions to standard STIT

We have noted above, in section 3.3, that other agent nested STIT presents semantic problems. We address these problems by reading the strict STIT as *influences*. Recall that influence does not necessarily guarantee that an agent will be able to bring a proposition, $\neg A$, about. An absolute guarantee would bring the statement under the remit of standard STIT. If an agent has all the pieces required for a particular influence then it will be in a position to guarantee $\neg A$. The pieces of an agent's influence may be dependent on other agents and this is where influence allows a semantically sound reading of nested other agent sentences. If α does something that allows β to bring about $\neg A$ then saying that α *influences* β which, in turn, *influences* A makes sense. We shall address operators and syntax in the following chapter where we develop a practical theory of influence. Substituting this into a nested STIT sentence, $[\alpha\ influences\colon [\beta\ influences\colon A]]$, is semantically sound with no cost to individual or group agency.

3.5 A first validation of nested influence

Our first exploration of nested influence question is set in a small world which has four cells arranged as a square and is populated by two agents. The agents, let's call them α and β, have different abilities, agent α

can give a token to β and β can use this one time token in any of the world's four cells. Each cell contains a seed which may randomly cause A to hold. The world operates in action cycles and when A is randomly brought about if the proposition has a finite (and random) life, if it is not countered – by agent β using a token – within this life then it will clear itself.

The purpose of this simple experiment is twofold, firstly it is to attempt to formally identify events which increase agent influence on the world by bringing about states that an agent is unable to bring about on its own and, secondly, an attempt at synthesising agent behaviour patterns in a manner that will allow us to have agents repeat behaviours leading to a greater incidence of such states.

3.5.1 Characterising influence

The idea of *influence extending behaviour* was introduced in section 3.4.3, we formalize the notion here and we also consider some potential difficulties. Informally, α giving a token to β is a necessary step which extends *β's* ability but is not, on its own, sufficient to clear an instance of A. *β's* using a token on A is necessary and sufficient (just in case A does not spontaneously clear) to counter A. The relationship between *influences* and STIT can, informally, be seen here. Necessity and sufficiency equates to stit whereas necessity alone equates to influence. A strict reading of $[\alpha\ stit\colon [\beta\ stit\colon \neg A]]$ is not consistent but by considering stit as encapsulating influence and by examining events that have occurred we are able adopt a consistent reading for statements of this form.

We characterize the extended influence part of ability as a variation in the mapping of histories to the choice partition seen by an agent at a moment. In figure 3.5 the *use token* choice partition at m_2 (given that A holds) has one certain future whereas the *use token* choice at m_3 is indeterminate. We extend the $Choice^m_\alpha$ notation of figure 2.4 in section 2.7.2 by prepending an h to indicate that some history is "choosable". That is that this history may be guaranteed by an agent choice. Thus, $h.Choice^m_\alpha$ reads some history, h is available to agent α at moment m by *α's* executing one of its choices. This allows us to consider the existence of a single "choosable" history when comparing choice sets. This extension permits us to represent a collocated action where one agent enables another and in doing so extends its ability in equations 3.1 and 3.2 indicating that a joint action has an effect of the distribution of histories causing it to differ from the distribution in an individual agents choice partitioning.

$$\exists h.Choice^m_{\alpha\|\beta}(h) \subsetneq Choice^m_\alpha(h) \tag{3.1}$$

or

$$\exists h.Choice^m_{\alpha\|\beta}(h) \subsetneq Choice^m_\beta(h) \tag{3.2}$$

Informally, joint action by α and β will alter the choice partitioning visible to α and β individually and will do so by either adding or removing histories or by refining the extant partitioning. These equations hold in figure 3.5 where a joint action carried out at I_0 leads to β being able to bring about A at the completion instant, I_0. The example of figure 3.5 is "short" in that the completion instant immediately follows the instant where β acquires the ability to see to it that A holds. Following the influence extending joint action β*'s* reaching a completion instant is contingent on it not doing anything which compromises that ability and this is something which we shall address in the following section.

Sequential action is, intuitively, rather different from cooperative actions. Agent influence in a collocated and cooperative action may be thought of as being commutative, α cooperating with β is the same as β cooperating with α. Sequential influence is not necessarily commutative. A γ type agent dropping a token in a cell *after* a δ type agent has visited that cell will not provide an enabling instant (looking at a closed sequence of events that have already occurred so that there is no possibility of δ returning) whereas a γ dropping a package in a cell *before* a δ visits will provide an enabling instant.

$$\exists h.Choice^m_{\gamma:\delta}(h) \subsetneq Choice^m_\delta(h) \tag{3.3}$$

Informally this reads that the choices available to δ at moment m when preceded by γ's action is contained within what δ's choices would have been in the absence of γ's action. Intuitively, γ's action plays a part in refining the distribution of histories in δ's choice partitioning so as to either remove uncertainty or add new histories and extend δ's ability.

We continue by exploring this informal characterisation of influence by exploring history traces produced by a simple experiment.

3.5.2 Examining initial experimental data

The dismissal of [α *stit*: [β *stit*: A]] as being impossible (section 3.3) is certainly true if one is viewing current events and considering the future. If α gives a token to β then it is not correct for us to say that α sees to it that β sees to it that A does not hold. If, however, we look at events that have happened and view these events in terms of agent ability we can say that α giving a token to β has seen to it that β countered A because α's action enabled β to do so.

Some agent histories are examined below, these are taken from data generated by simple experimental systems. The leftmost column indicates the achieved state following the agent actions detailed in the two right hand columns. The agent state data in these columns are those before agent action and are, thus, percepts from the immediately preceding state. We term such percepts preceding an agent's choice *precepts* and, similarly percepts following an agent's choice are termed *postcepts*.

Definition 19 *Given an agent situated in an environment, which it is able to perceive and manipulate, we define precepts as an agent's set of percepts before that agent makes a choice that may alter its environment.*

Definition 20 *Given an agent situated in an environment, which it is able to perceive and manipulate, we define postcepts as an agent's set of percepts after that agent has made a choice that may have altered its environment.*

If, for example, α gives a token to β then this will not show up on the agent state table until the next cycle. Any transitions indicated are those between the previous and current cycle. This is illustrated in table 3.1 which lists a series of events in a 2 x 2 cell agent world.

Table 3.1: Example history fragment table

State ID	Transition	α	β
16		S	S
15		EA	E
14		W	N
13		G	W
12		S	E

Agent actions and state are represented using the following code; G = attempted to pass token to another agent, T = used token, X = No action, A = this agent perceives A (at start of cycle), N,W,S,E = moved in specified direction. Additionally, α and β mean that the agent sees the other agent in its current cell. Where cell numbers are mentioned these are zero based and run from bottom left to top right. We simply state *Cleared* where a A changes to $\neg A$ and make no distinction between forcible clearing and or the proposition simply coming to the end of its life. When β is carrying an enabling token this is indicated, for convenience, by an asterisk in tables.

Each fragment is outlined in two ways, the transition column indicates a global or observer's view by explicitly indicating proposition transitions. The α and β columns provide a situated coaching agent view detailing the agent level data that coaching agents will work with. Note that it is assumed that coaches know that propositions do not move so when a A drops from an agent's percept data after a move this does not necessarily mean that A has been countered or has come to the end of its life.

3.5.3 Experimental data – fragment one

A number of experiments were run on a simple simulation of this four cell, two agent world. The tables in this and the following sections represent data gathered from the log files generated by these experiments. This first fragment (illustrated by figure 8.2 in appendix 2) represents three possible routes from enabling –

when α gives a token to β – to a proposition being countered. These sequences range from two steps long to five steps and each begins with α and β meeting in cell 1 having moved south and east respectively in the previous cycle. The sequence of fragment 1.1, table 3.2 starts at the bottom of the table (the top entry

Table 3.2: Fragment 1.1, worlds 12–16, $\neg A$ state 16.

State ID	Transition	α	β
16	Cleared	AS	U
15		EA	EA*
14		W	N*
13	Enabling	GA	W*
12		SA	EA

being the most recent) with agents α and β moving south and east, respectively, into cell 1. Both agents detect that A holds in this cell. α gives a token to β and β, now carrying the token, moves west to cell 0, an empty cell. Step 14 is a simple move step with both agents moving into other empty cells. Step 15 sees both return to cells where A holds. In the following step α attempts a south move which, because of its position in the south row of its world is effectively, a null move whilst β uses its token and counters the instance of A in that cell. We represent this series of moves by temporarily using slightly modified stit notation with world indices added for convenience to allow cross reference to the data that are in tables. Thus we have $[\alpha\ stit_{w13} : [\beta\ stit_{w16} : A]]$ meaning that α sees to it at world 13 that β is able to see to it that A does not hold at world 16.

How can we view what is happening here? We bound the sequence and declaring two sets of instants, I_w – the set containing the witnessing moment – at the start of the sequence and I_c – the set containing the completion moment where an agent makes the choice that brings about $\neg A$. This is illustrated in figure 3.6. Our intuition is that, in this case, the sequence begins at m_0 with α giving a token to β. In this scenario β's ability to see to it that $\neg A$ at some point in the future is contingent on α giving a token to β and, if there is no instance of A at the agent's joint location, β not immediately using the token.

There may be an indeterminate number of instants between I_w and I_c, this may introduce uncertainty about result of the sequence. This uncertainty may be handled by reading the expression, as noted already,

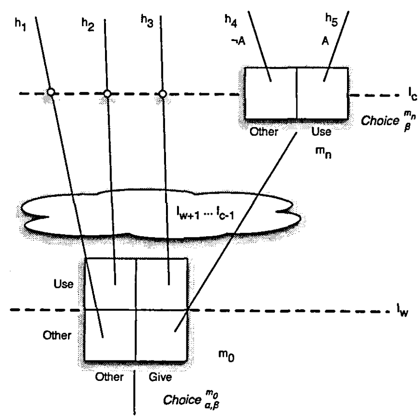


Figure 3.6: Fragment one from experimental data, does [α *stit*: [β *stit*: A]]?

as α sees to it that β is able to see to it that A. The truth of the statement is contingent on *β's* agency. The indeterminacy of the number of moments is computationally difficult and we address this problem later when we examine details of the coaching agent in section 6.6.1.

Our intuition is that any semantics used must have a negative condition (which holds until the completion instant), there is some uncertainty between the witnessing moment and the closing moment. Uncertainty, at its simplest, means that there is a chance of A not happening. This is perhaps best thought of as there being a possibility of β losing its ability to see to it that A by simply using its token in a cell where $\neg A$ already holds. This leads to an interesting question – is *refraining from acting* acting? We avoid this issue by considering the *positive commission* aspect of actions thus β not using a token may be thought of as β acting in such a way as to maintain its potential influence.

3.5.4 Experimental data – fragment two

Fragment one, discussed above and illustrated in figure 8.2, contained three history paths from world 12 and each of the three paths led to a state where an instance of A was countered. Fragment two (illustrated in appendix 2 as figure 8.3) contains two possible paths and only one of these leads to an instance of A being countered. As in fragment one the interesting part of the sequence starts at world 21 where, as in world 12, α passes a token to β and β does not use it immediately.

Table 3.3: Fragment 2.1, worlds 20–21/26–28, A countered at state 27.

State ID	Transition	α	β
28		W	NA
27	Cleared	EA	U
26		SA	SA*
21	Enabling	G	EA*
20		EA	NA

Table 3.4: Fragment 2.2, worlds 20–25, A not countered.

State ID	Transition	α	β
25		NA	UA
24		EA	EA
23	Disabling	W	U
22		SA	W*
21	Enabling	G	EA*
20		EA	NA

Fragments 2.1 and 2.2 are detailed in tables 3.3 and 3.4. By compressing the two evolutions and using the same layout as figure 3.6 we get the scenario illustrated in figure 3.7. The witnessing moment is at world 21, at world 27 agent 1 uses its token in the same cell as an instance of A and counters it but at world 25 using the token does nothing. Using the token at m_c is inconsistent despite the previous enabling action at m_w. Clearly something is happening at an instant between I_w and I_c to disable the agent. Our privileged view and knowledge of the world tells us that it is using the token at world 23 that removes the agent's ability. In the cloud of uncertainty between I_w and β's use action at I_c when A does not hold removes the agent's ability to see to it that A and closes the sentence, illustrated in figure 3.8. On reflection this is not really any different from the illustration in figure 3.6

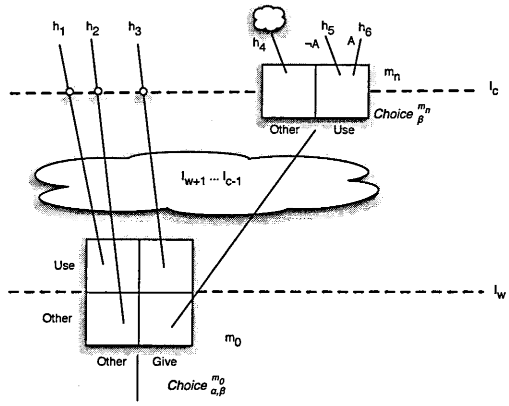


Figure 3.7: Fragment two from experimental data representing [α *stit*: [β *stit*: A]] $\wedge \neg$[α *stit*: [β *stit*: A]]

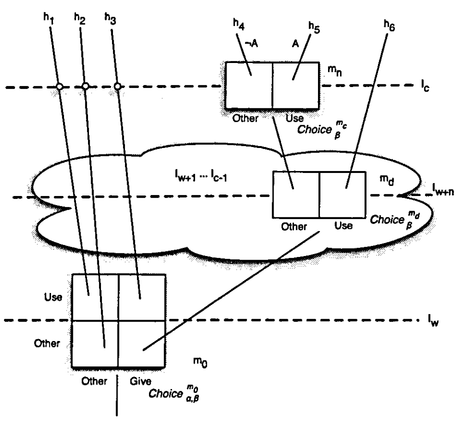


Figure 3.8: Fragment one from experimental data with intermediate choice for β

3.5.5 Identifying extended influence from observations

Extended influence, intuitively, manifests itself in settings where more than one agent is needed to complete a behaviour. Each agent has an influence on the execution of the action. We have noted above, in section 3.5.1, that joint action, $\alpha \| \beta$ may be expressed or identified by comparing choice partitioning. Joint action exists if α and β *together* have access to a history that would not have been available had they been operating individually.

$$\exists h.Choice^m_{\alpha\|\beta}(h) \subsetneq Choice^m_\alpha$$

or

$$\exists h.Choice^m_{\alpha\|\beta}(h) \subsetneq Choice^m_\beta$$

This is evident in figures 3.6, 3.7 and 3.8 where, in each case, when a joint action is carried out at I_w leads to β being able to extinguish a fire at the completion instant, I_c.

3.5.6 Noisy influence and the lifespan of influence

The examples illustrated in figures 3.6, 3.7 and 3.8 are complex in that there is uncertainty between the enabling and completion instants. Let us take a very simple case, worlds 12/17–18 from history one, and remove the cloud of uncertainty from the illustration. The key point in the 12/17–18 example is that the witnessing instant is immediately followed by the completion instant, this is shown in tables 3.5 and 3.6, note that the witnessing instant has been labelled I_0 and the completion instant I_c.

Table 3.5: Fragment 1.2, 12 / 17–18, *A* countered at state 18.

State ID	Transition	α	β
18	Cleared	NA	T
17		GA	EA*
12		SA	EA

Table 3.6: Fragment 1.3, 12 / 17–18, A countered at state 18.

State ID	Transition	α	β
19	Cleared	G	T
12		SA	EA

The joint agent action is illustrated in figure 3.9, without the cloud of uncertainty we can see that the give token / do not use token combination at the witnessing instant delivers agent β directly to the completion instant and world 18 where β has a choice that allows it to see to it that an instance of A is countered.

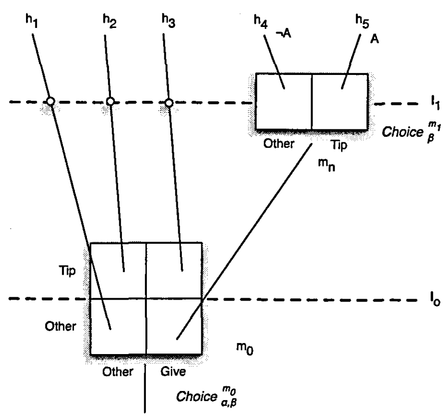


Figure 3.9: Fragment one from experimental data with uncertainty removed, does [α *stit*: [β *stit*: A]]?

If we consider agents α and β individually, figures 3.10 and 3.11 respectively. In order for β to reach world 18 it is necessary that β be in the same location as an instance of A and for β to be carrying a token. α is unable to counter A on its own so, clearly, in the absence of β α is unable to reach the choice at world 18 (which is a member of I_c).

If β is acting on its own and β arrives at the witnessing moment not carrying a token then, like α, β has no choice available to it that will bring it to world 18 at the next instant.

Intuitively one stit like construct handles the entry to and another the exit from a sequence of events that brings about A being countered. If the entry stit like construct is not satisfied then the agents do not enter into a state that may lead to an instance of A being countered.

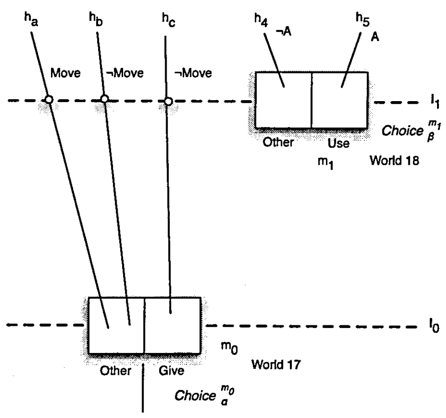


Figure 3.10: Fragment one from experimental data representing α's situated perspective

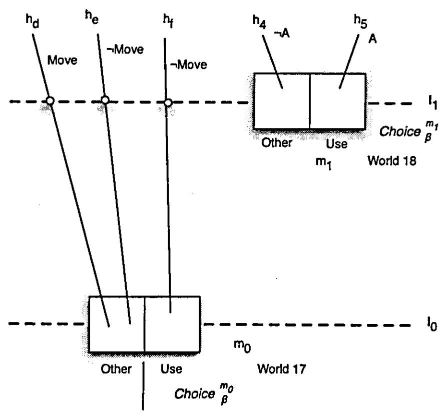


Figure 3.11: Fragment one from experimental data representing β's situated perspective

3.6 Summary - an outline theory of influence

We have introduced the notion of influence and explored its potential as an agent ability operator that allows a meaningful interpretation of other agent nested STIT type constructs. We have extended the notion of influence to account for what we term serial and parallel influence. Our consideration of influence has been in the abstract and this indicates that influence has an advantage over STIT in that it admits cases where agent ability is dependent on factors outside of its direct control, without entailing any additional complexity over the extant STIT like construct. Our intention is to apply our theory in a practical setting, a consideration of *noisy influence* is a first step in this direction. The ability of the theory to carry a potential or undischarged influence allows agents to operate in a noisy environment. In the following chapter we take our notions of influence and apply them to a practical setting in a noisy environment. In this setting we attempt to *identify* where influence occurs and to do so purely on the basis of observations so as to allow us to develop a practical theory of influence.

Chapter 4

Developing a practical theory of influence

We have extended the semantic reach of nested STIT constructs by introducing the notion of influence as an alternative to strict STIT. The previous chapter built an outline to a theory of influence, in this chapter we fill in that outline and work towards a practical theory of influence. We begin by considering an algebra for agents and then proceed to develop a theory which allows us to apply influence in this setting. The development of this practical theory and the description of agent algebra will then put us in a position to consider, in the following chapter, how coaching agents operate and how they may manage any hypotheses formed from their observations of agent behaviour. Our investigation of the relationship between STIT and influence will follow the path illustrated in figure 4.1.

4.1 Preliminaries, discrete branching time and an instantaneous STIT

Standard branching time is *dense* in that its ordering for moments is simply an ordering and does not contain absolute time interval data. We need to have a guaranteed minimum time interval between instants so as to guarantee both an *intra-instant* interval where an agent may complete an operation cycle and that the results of agent actions will be evident at the immediately following instant. In the standard branching time model if $m_1 < m_2$ we know that m_1 occurs before m_2 but we do not know by how much or even if m_2 is the moment immediately following m_1. This presents problems for a bounded reactive agent system which,

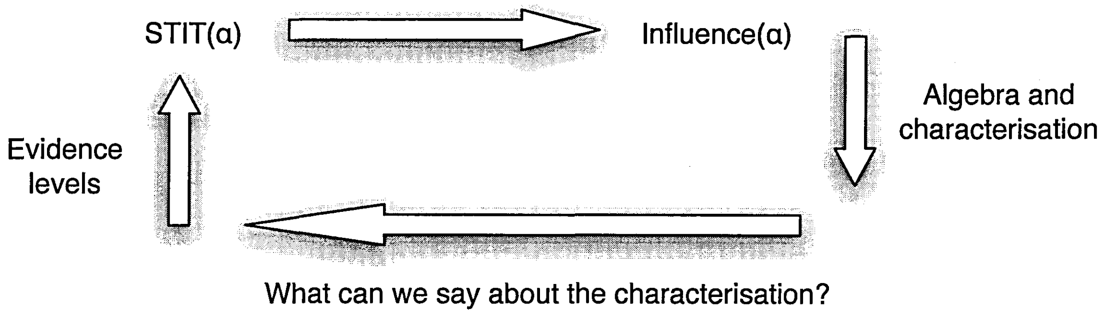


Figure 4.1: The investigation path from STIT to influence and back

as Chang et al. [28] note, is in constant interaction with its environment. How many instants lie between two given instants? Tulenheimo [115] indicates that there is a need for investigation into time division for tree like structures. One way of dealing with this may be to introduce a *next instant* operator. Broersen [22] describes a *next–time* relation but this is related more to the serial and deterministic nature of an agent's actions than the absolute time involved. Broersen also describes a STIT extension to alternating time temporal logic but this, again, lacks a deterministic representation of time. The notion of a discrete time STIT based on a temporal next operator, X, (equation 4.1) is described in [23]. Time is discrete in the sense that a following moment or instant is guaranteed to be the next in a sequence. This is close to what we need but there is still no guaranteed time interval between moments or instants and, thus, no means of describing or guaranteeing an instantaneous action.

$$\mathcal{M}, m/h \models \mathrm{X}_{\varphi} \Longleftrightarrow \exists m' \in h(m < m', \mathcal{M}, m'/h \models \varphi, \nexists m'' \in h(m < m'' < m')) \tag{4.1}$$

We adopt a simplistic approach to this and assume a discrete background time and impose its ordering on instants.

Definition 21 *Given a set of time ticks, t and the natural numbers* $\mathbb{N}$*. Let time* $= \mathbb{N}^+$ *We refer to time as t with* t_0 *being the first time tick and* $t_0 < t_1 < t_2 \ldots$ *with no intervening ticks. Ticks are separated by an arbitrary but consistent time interval.*

We extend our definition of instants, 11, by adding time indices.

Definition 22 *Given a branching time tree, Tree containing a set of moments, m, and histories, h and a set of discrete time ticks, t. An instant is a horizontal partitioning of a branching time tree which groups events occurring simultaneously. Instants are linearly ordered and lie on time ticks which may be used as an index to uniquely identify an instant within a context. Instant i_n, for example lies on time tick t_n.*

This allows us to define when one instant follows another:

Definition 23 *Given two instants, i_a and i_b, instant i_b, follows instant i_a iff $t_b > t_a$.*

and we may also say when one instant *immediately* follows another:

Definition 24 *Given two instants, i_a and i_b, instant i_b, immediately follows instant i_a iff $t_b > t_a$ and $\nexists t_c$ such that $t_b > t_c > t_a$.*

We do not specify a time interval between ticks but note that each instant notionally contains an agent cycle allowing agents to acquire new behaviour patterns if available, gather precepts, select and execute an action, gather postcepts and drop historical data. This is an agent-internal behaviour cycle that we shall meet again so we state this as a definition for agents in this work.

Definition 25 *Given an agent α which may perceive and act in its environment. An operating cycle for α consists of gathering a set of pre action percepts, precepts, selecting and executing an action then gathering a section of post action percepts, postcepts. These steps are all contained within the discrete time of a single instant.*

Further, any changes in the environment caused by that agent's action will stabilise and be visible to precepts at the next instant.

This relationship between instants and a background discrete time framework allows us to entertain meaningful notions of instantaneous actions and to define an instantaneous STIT evaluation rule for such events.

Where α is an agent and m/h is an index from a stit model $\mathcal{M}$:

$$\begin{array}{rcll} \mathcal{M}, m/h & \models & [\alpha\ istit\colon A] \text{ iff} & (4.2) \\ & (i) & Choice^{m}_{\alpha}(h) \subseteq |A|^{\mathcal{M}}_{m} \text{ and } |A|^{\mathcal{M}}_{m} \neq H_m & \\ & (ii) & \exists I \text{ at time } t \text{ such that } m \in I, \exists I' \text{ at time } t+1 & \\ & (iii) & \forall h \in Choice^{m}_{\alpha}(h), \exists m' \in h \text{ and } m' \in I' \text{ where } A \text{ holds.} & \end{array}$$

Equation 4.2(i) is the standard deliberative stit rule which states a positive condition, that A be true on all histories in a partition available at moment m, $Choice^{m}_{\alpha}(h) \subseteq |A|^{\mathcal{M}}_{m}$ and a negative condition that the set of histories where A holds is not the full set of possible histories branching forwards from m. Equation 4.2(ii) states that there are two instants, I and I', separated by one time tick and that moment , m, belongs to the instant immediately preceding I'. Equation 4.2(iii) is, perhaps, the most awkward. When an agent is acting in such a way as to influence another then the results or effects of that action must be available to other agents. We state this by saying that there are moments belonging to I' and lying on all of the histories satisfying A from a choice partition where A holds *at that moment*.

This notion of an instantaneous STIT is a convenience which gives us an implicit synchronisation mechanism. This allows us to assume that in a sequence where one agent's ability depends on the actions of a previous agent then the results of the previous action will be available at the start of an instant.

4.2 Requirements for an agent algebra

We intend to apply our notion of influence to purely reactive agents and this brings some difficulties. Earlier work such as Ferber's approach to influence allows for one agent *forcing* a decision for other agents (see Michel [86]) and this strips agency from some of the parties involved in a joint action. Clearly preservation of agency requires that each party involved in a joint action have some form of synchronisation that preserves individual agency. This seems like an opportunity for agent communication but as we indicated earlier there is a trade off between agent and communications complexity (see Bryson [27]) and we have elected to use implicit communications driven by agent percepts and an implicit expectation of other agent behaviour. When one agent meets another agent and they are able to jointly extend their individual influence then what do they do? van der Hoek at al. [116] propose a counterfactual extension to a temporal logic. This is a ternary operator which assigns a value to any suppositions that an agent may need to base its decisions on, these may be questions like *will the other agent co-operate or not.* We have briefly mentioned, in section 2.9.1, a ternary approach to belief and this is something which we shall revisit when considering coaching agent operation. Such reasoning is beyond our simple reactive agents so we adopt a default logic approach. If an agent has a behaviour pattern that is contingent on the presence of another agent and the conditions for that behaviour are satisfied then the agent will simply go ahead and follow that behaviour pattern. The agent's action selection will never be absolutely guaranteed because behaviour patterns are biases towards rather than prescriptions for a particular action. Our approach to describing agent behaviour and, consequently, the algebra of their behaviour is based on this default approach.

4.3 An agent algebra and observable behaviours

We begin our consideration of an agent algebra by recalling our touchstone observation by Milner [87].

> The behaviour of a system is exactly what is observable.

We have considered a single agent's influence on some aspect of its environment and noted that there are sets of circumstances where a *single* agent may have no influence but where a *group* of agents may. In order to reason about the behaviour of groups of agents we need some form of representation, an algebra which allows us to manipulate agents so as to build group behaviours. There are many variations of process calculi based the notion of communicating sequential parallel processes, such as those by Milner [87] and Hoare [62], but these may not be best suited for our needs since, as per observations 1 and 2, agents are independent, autonomous entities.

We have stated our intention to maintain the independence of agents throughout the development of our theory. Standard process algebraic approaches allow for processes to be combined so as to construct new processes – take P_1, chain it with P_2 to get a new process P_3 – the processes have no say in the matter. Combining agents and maintaining agency seem to be mutually exclusive operation. If α *chooses* to act in a certain way and β *chooses* to act in a complementary way then together the agents may achieve something that they could not have done individually. If α and β are forced to act in a certain way then they lose some of their agency. We begin with this notion of *choice* and use this as a grounding for modelling our agents as state machines.

4.3.1 Defining agent choices

We assume an autonomous agent which may choose how to behave from a number of possible choices.

Definition 26 *For an agent α with a set of choices $Choice_\alpha$, and some choice $K \in Choice_\alpha$ we will refer to the choice of K by α at some moment as α/K.*

When an agent makes a choice it is carrying out some action which, when viewed against the background branching time framework, selects a set of possible future histories lying within the equivalence class defined by the histories branching forwards from that particular choice. Choices and actions are closely related but not necessarily the same because an agent that is not suitably equipped may execute a choice but be unable to complete the action associated with that choice. We are investigating the influence of an agent and that agent's choices on some aspect of the agent's environment in order to evaluate whether or not α/K has influence over A. α may have a number of choices in addition to K but since we are concerned only with the influence of α/K over A we use a simplified characterisation of the agents choices.

Definition 27 *For an agent, α, with choices, $Choice_\alpha$, and for a choice element $K \in Choice_\alpha$ we define the negation of K, $\neg K$ as the set $Choice_\alpha \backslash K$.*

If, for example, an agent has the set of choices $\{J, K, L, M\}$ then the set $\neg K$ will be $\{J, L, M\}$. When considering the influence of α/K over A we will do so with a choice set $\{\neg K, K\}$ with $\neg K$ representing any agent choice other than K. It may be that some other choice available to α will bring about A but this is of no concern to us as we are concerned only with the influence of α/K over A, if α makes a different choice, say L, that influences A then that will be the influence of α/L over A.

4.3.2 An algebra based on choices

If we consider agent interaction as having two dimensions, events and agents, then the more major of these is events and these events are directly related to agent choice, Our means of representing this is to write agent driven events as choices subscripted by agents. For example, parallel agent action may be written as $K_\alpha || L_\beta$. This extends readily to groups of events, for a set of choices $C = \{K, L\}$ and a set of agents, $G = \{\alpha, \beta\}$ we may say $C_G||/A$.

Similarly the serial agent action may be written as $K_\alpha\ ;\ L_\beta$. Shifting to a set or group based notation is more complex because serial behaviour requires some form of ordering. We assume that this is implicit in the ordering of the actions. The agent ordering is not strict, what is important is that for each event in C there is at least one agent in the set of agents that is capable of making the required choice in appropriate circumstances.

The set of events required to bring about A and, in the serial case, the sequencing of those events may be fixed but this is not necessarily the case for the set of agents required to drive those events. Extending the serial case to a more general notation. This extends readily to groups of events, for a set of choices $C = \{K_2, L_1\}$ and a set of agents, $G = \{\alpha, \beta\}$ we may say $C_G\ ;\ /A$ and we note that the ordering on the group of choices C is a strict ordering which does not necessarily apply to the group of agents G.

This group of actions postfixed by a serial or parallel operator allows us to combine a number of aggregate behaviours. For example, a serial action followed by a parallel action may be represented as $C_{G_1}\ ;\ D_{G_2}||/A$

4.4 Agent empirical studies of influence

Observation plays a great part in our analysis of agent behaviour. We have adopted a privileged observers role in earlier discussions and have indicated that this is unsuitable for coaching agents. Coaching agents will observe how other agents in the world behave from a similarly situated perspective. We do, however, give coaching agents an insight into agent operation by allowing them to see what choices an agent has made at any particular moment. Coaching agents operate with a family of hypotheses. These form a sequence based on the hypothesis that an agent has influence over its environment. If the coaching agent observes evidence of influence of some sort and that evidence is unable to satisfy a single agent behaviour hypothesis then the coach will test the evidence against two-agent hypotheses that agents may influence their environment jointly. The intuition behind the second of these hypotheses is that agents may act in ways that *extend* their influence over what they may achieve individually. Two agents may carry a larger object than one or a bucket brigade may transfer objects more quickly than agents working individually.

Other researchers have introduced logics for dealing with evidence. Halpern and Pucella [56], for example, describe a logic for reasoning about evidence. This approach differs fundamentally from ours as it regards evidence as a function from prior beliefs to posterior beliefs. Prior beliefs are those held before making observations and posterior beliefs are those held afterwards. Although our system may be thought of as holding beliefs these are simple and the coaching agents do not hold any prior beliefs because their operation is built on observation and observations provide evidence to generate and rank hypotheses.

4.4.1 From STIT to influence, leads to and may lead to operators

Intuitively if some agent choice pair, α/K, has *influence* over A we would expect that A should hold following α/K. We will talk of *A following α/K* so to avoid ambiguity we state:

Definition 28 *Given an agent α acting at moment m with $m \in$ instant I by choosing K from its available choices brings about A then we say that A* follows *α/K when A evaluates true at the immediately following instant, $I + 1$.*

It may occasionally be more convenient to talk of an agent/action pair immediately *leading to A* so, again to avoid ambiguity, we introduce an leads to operator:

Definition 29 *Given an agent α acting at moment $m \in$ some instant I by choosing K from its available choices then we say either that α/K leads to A and write this as $\alpha/K \rightsquigarrow A$ when A evaluates true at the immediately following instant, $I + 1$.*

$\alpha/K \rightsquigarrow A$ is a necessary part of demonstrating that α/K, has *influence* over A but on its own it is not sufficient. If A were constant and α had no influence then not only would A follow α/K, but it would also follow $\alpha/\neg K$. We require that A is not invariable in order for α to have influence and we expect to see instances of $\neg A$ following $\alpha/\neg K$ as evidence of α's ability to influence A. If $\neg A$ follows α/K then it would appear that α/K has no influence over A.

Coaching agents will gather evidence, by observation and by gathering details of agent choices, to support or deny their hypotheses of agent influence. We outline three classes of observed evidence, positive evidence, negative evidence and counter evidence, by the following observations:

Observation 7 *Positive evidence for a hypothesis that α/K has influence over A is an observation of A following α/K.*

Observation 8 *Negative evidence supporting the hypothesis that α/K has influence is an observation of $\neg A$ following $\alpha/\neg K$.*

Observation 9 *Counter evidence for the influence of α/K over A is an observation of $\neg A$ following α/K.*

We note that these look rather similar to the ternary belief operator outlined in section 2.9.1.

The negative evidence requirement of $\neg A$ following $\alpha/\neg K$ need not be consistently observed for a number of reasons. Recall that, by definition 27, $\neg K$ is the complement of K and $Choice_\alpha$ and that another choice by α may also have influence over A without compromising the influence of α/K over A. It may also be that a proposition already holds, if α gives β a token when β already has a token then β's token count does not increase. Similarly if α does not give a token holding β a token then β remains as a token holder This second case allows us to define a fourth class of evidence, if A follows $\alpha/\neg K$ then we can infer neither influence nor lack of influence and make the following observation:

Observation 10 *An observation of A following $\alpha/\neg K$ is neutral and gives no evidence for the influence of α/K over A.*

This brings us to a point where we may refine our broad view of agent influence and restate this as the first of the coaching agent's hypotheses. The coaching agent's hypotheses are intended to allow it to interpret observed behaviour in a manner that corresponds with an agent's choice partitioning and the possible futures that may result from agent choices.

Hypothesis 1 *The Single agent influence hypothesis is that α/K has influence over A. Evidence for this hypothesis will be:*

i) A following α/K

ii) At least one $\neg A$ following $\alpha/\neg K$

A counter hypothesis, a hypothesis that α/K has *no influence* over A, is indicated by $\neg A$ following α/K.

Hypothesis 2 *The Single agent influence null hypothesis is that α/K has no influence over A. Evidence for this will be:*

i) $\neg A$ following α/K

We now have a hypothesis and null hypothesis which allow us to identify single agent influence. These are supported by a partitioning of observed evidence into classes which either support, deny or are neutral towards the hypothesis. So far, these hypotheses cover only the single agent case. We shall investigate this in a more depth then extend the single agent hypotheses to deal with two agent influence. After considering the two agent influence hypotheses we will have sufficient information to be able to consider the general case and generate hypotheses for communities of agents.

4.5 Single agent influence at a single moment

The single agent hypothesis that α/K has influence over A is examined against a simple single agent example. Figure 4.2 illustrates all of the possible outcomes for a single proposition, A, at a moment contained in an instant, I. By examination of the possible histories, $\{h_1, h_2, h_3, h_4\}$ and evaluating A following α's choice we see that:

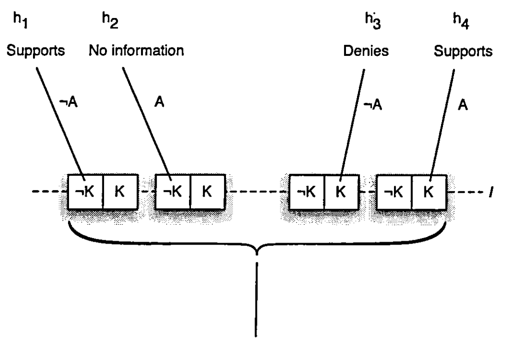


Figure 4.2: Evidence classification for Single agent influence on A

- History h_1 shows $\neg A$ following $\alpha/\neg K$ and this falls into our class of negative evidence. It also partially supports hypothesis 1 by providing evidence that $\neg A$ is possible and that it may follow $\alpha/\neg K$.
- History h_2, where A follows $\alpha/\neg K$ lies in the no evidence class and provides no information for the hypothesis.
- History h_3 indicates $\neg A$ following α/K. This falls into the counter evidence class of observation 8 and satisfies the null hypothesis indicating that, on this history, α/K has no influence over A.
- History h_4 indicates where A follows α/K satisfies the influence hypothesis and falls into the positive evidence class.

For this simple case we see that each of the possible histories may fall into only one of our evidence classes which removes the possibility of ambiguous evidence. Ignoring h_2 which provides no evidence, we see that each history may satisfy either the hypothesis or the null hypothesis but not both. For the simple case of a single agent at a single moment the single agent hypothesis and evidence classification hold.

Let us consider this further with some simple examples where we set an agent in an example world in order to consider simultaneous events. Assume that α has a fire extinguisher and is standing beside a piece of faulty electrical equipment which is not burning but may spontaneously ignite. We read A as *equipment not burning* and $\neg A$ as *equipment burning*. If α uses the fire extinguisher, K, at a moment m belonging to some instant I then there are three possible states for A. A may hold so that the equipment is not burning at I or it may not hold so that the equipment is burning at I. We call these *case A* and *case B* respectively. The third state, *case C* is a little more complex, consider the equipment spontaneously igniting at I just as α chooses to and uses its fire extinguisher. Recall that we cast this system against a background of discrete, branching time with time steps bounded by *instants* and agents acting at these instants. We informally adopt an instantaneous view of actions which means that an action at one instant is evaluated at the next instant. This instantaneous view is described more formally in section 5.1 at the end of this chapter. A fire which is fizzling on the boundary of ignition may gradually take hold so its start may be across a number of instants. If α uses a fire extinguisher at I and that fire extinguisher works then it will destroy the environment required for supporting a fire. By the time the evaluation instant arrives the extinguisher will have extinguished the weak fire. It is not unreasonable to assume – against a framework of discrete branching time – that in such circumstances a fire starting at the same instant as an extinguisher is used will not continue to burn. Case C, there, is similar to that of case B and we assume a starting fire is extant at I. In case A, if α chooses K then we will observe an h_4 pattern and if α chooses $\neg K$ then we will observe an h_2 pattern. Case B will give h_4 and h_1 patterns for K and $\neg K$ respectively and case C will also give h_4 and h_1 patterns. For a simple, single agent example where, intuitively, the agent has influence over A the hypothesis and evidence classes hold. We summarise these histories in table 4.1.

Table 4.1: Single agent influence

	K	A	$\alpha/K \rightsquigarrow A$
1	$\neg K$	$\neg A$	Supports
2	$\neg K$	A	Neutral
3	K	$\neg A$	Denies
4	K	A	Supports

The balance of evidence that a coaching agent observes contains some subtle information. The fourth row of table 4.1, where A follows K supports the hypothesis and the third row, $\neg A$ following K. denies it. There is nothing contentious in this interpretation. The second row is neutral, if an agent knows that K

leads to A and A already holds then why K? The cases represented in the first and third rows carry more information than merely supporting and denying the hypothesis. If a coaching agent sees neither of these then this indicates that $\neg A$ has not occurred whenever α has acted then this indicates that A may well be invariant, we have evidence that A follows α/K we have no evidence that $\neg A$ is possible. In order for the coaching agent to assume that α has influence over A it must have evidence that $\neg A$ is possible and that $\neg A$ follows $\alpha/\neg K$, as described in observation 8, at least once.

4.6 Joint multiple agent influence at a single moment

Consider the above scenario with a slight change. Instead of a self contained fire extinguisher α has a fire hose. However, the hose is not as self contained as the fire extinguisher, it requires a tender at the other end and this tender must be manned. This manning requirement may be that another agent, β needs to continuously operate a pump or that β needs only to connect the hose and open the associated valve.

If β were to fail to do either of these then α would be stripped of ability and we would see only h_3 pattern histories. If β occasionally pumps foam or occasionally opens the valve then we will occasionally observe h_3 and h_4 pattern histories where there were always h_4 patterns before. The incidence of h_1 and h_2 histories will be unchanged. α will appear to have influence occasionally. This leads us to our second hypothesis that other agents may contribute to α*'s* influence. With a hose and tender it seems that α alone has no influence over A since it needs assistance from β. β alone is unable to influence A because it needs α to aim and use the hose. We call such influence examples *extended influence* and two classes of this are illustrated by the hose example. If β needs to pump foam as α sprays it at the faulty electronics then both are contributing influence at the same time and in *parallel.* If β simply needs to open a valve before α uses the hose then β*'s* action must precede α*'s* using the hose and the two actions operate *serially.* We shall examine these *parallel* and *serial* behaviours in following sections we consider both of these joint action types to examine how they fit into our notion of influence and to characterise them in such a manner that a coaching agent may be able to observe and identify them.

If α/K does not consistently exhibit influence over A then the coaching agent may hypothesise that some other agent is occasionally assisting and extending the influence of both agents. There are two possible simple scenarios for this. The first is joint parallel influence where, for example, β must continue to pump foam as α uses the hose and joint serial influence which we consider later.

Before proceeding we extend the α/K notation of definition 26

Definition 30 *For a group of choices, $C = \{K, L, \ldots, N\}$, and a group of agents $G = \{\alpha, \beta, \ldots, \delta\}$ with $\{K \in C : \exists\alpha \in G : K \in Choice_\alpha\}$ and $|C| = |G|$ we will refer to the simultaneous choice of C by G (at some moment) as $G||/C$.*

We extend choice negation to the multi agent parallel case where it becomes more complex. Intuitively for a set of choices spanning multiple agents the negation of that set is a set where at least one agent choice is is not contained in the set of required choices.

Definition 31 *For a group of agents, G, a group of choice sets $Choice_G = \{Choice_\alpha, Choice_\beta, \ldots, Choice_\delta\}$ and a group of choices $C = \{K \in Choice_\alpha, L \in Choice_\beta, \ldots, N \in Choice_\delta\}$ we generate a negation of the choice sets $\neg Choice_G = \{\neg K, \neg L, \ldots, \neg N\}$ and a complement set of choices $C' = \{K' \in \neg K, L' \in \neg L, \ldots, N' \in \neg N\}$. The negation of C, $\neg C$ is a set of choices with at least one element $\in C'$.*

Moving on, we now consider parallel agent choice:

Definition 32 *Given a group of agents, G, and a set of choices C and an instant I we say that these agents have* parallel *influence over A if and only if $G||/C$ has influence over A.*

Coaching agents may look for one of the two types of two agent influence outlined above. For this to be the case then it must have been unable to find clear evidence supporting its single agent influence hypothesis. We consider the two agent parallel case before considering the two agent serial case in the following section.

Having been unable to observe evidence supporting the single agent influence hypothesis the coach may examine its observations against a two agent parallel hypothesis. There are now two agents involved in influencing A so the evidence types of observations $7 \ldots 10$ – which are predicated on single agent influence – will need to be extended.

We make the following observations for evidence of two agent parallel action:

Observation 11 *By positive evidence for the influence of* $\{\alpha,\beta\}||/\{K,L\}$ *over A, we mean an observation of A following* $\{\alpha,\beta\}||/\{K,L\}$.

Observation 12 *By negative evidence supporting the idea of the influence of* $\{\alpha,\beta\}||/\{K,L\}$ *over A, we mean an observation of* $\neg A$ *following* $\{\alpha,\beta\}||/\neg\{K,L\}$.

Observation 13 *By counter evidence for the influence of* $\{\alpha,\beta\}||/\{K,L\}$ *over A, we mean an observation of* $\neg A$ *following* $\{\alpha,\beta\}||/\{K,L\}$.

Positive, negative and counter evidence are simple extensions of the single agent case into the two agent domain. We also have a requirement that that the group of agents act at the same instant. Neutral evidence is also similar thus A following $\{\alpha,\beta\}||/\neg\{K,L\}$ provides no useful information. If, in the context of parallel actions, α uses its hose on the burning electronics, β is next to α and not at the hose tender then there may be some other agent, γ operating, the pump. A following $\{\alpha,\beta\}||/\neg\{K,L\}$, where $\neg\{K,L\}$ is $\{K,\neg L\}$, is simply the occasional positive influence observed for α in the single agent influence domain being cast into an inappropriate two agent domain. The other agent here is γ and the observation of A simply indicates that α occasionally has some influence over A and the evidence may be considered neutral. If $\neg A$ follows $\{\alpha,\beta\}||/\neg\{K,L\}$, where $\neg\{K,L\}$ is $\{K,\neg L\}$, then the other agent may not be operating the pump.

The extension of our view of neutral evidence into the two agent domain weakens certain single agent evidence observations so as to account for the additional uncertainties involved in reasoning about multiple entities.

Observation 14 *An observation of A following* $\{\alpha,\beta\}||/\neg\{K,L\}$ *gives no evidence supporting the notion of* $\{\alpha,\beta\}||/\{K,L\}$ *having influence over A.*

The two agent parallel action case is a simple extension of the single agent case considered above. We now state these observations as observed behaviour hypotheses.

Hypothesis 3 *The two agent parallel influence hypothesis is that* $\{\alpha,\beta\}||/\{K,L\}$ *has influence over A. Evidence for this hypothesis will be:*

i) A following $\{\alpha,\beta\}||/\{K,L\}$

ii) At least one $\neg A$ *following* $\{\alpha,\beta\}||/\neg\{K,L\}$

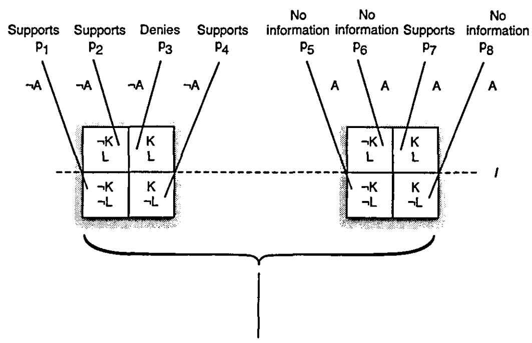


Figure 4.3: Evidence classification for parallel, two agent influence on A

iii) No evidence that either α or β has individual influence over A

A counter hypothesis, a hypothesis that $\{\alpha, \beta\}||/\{K, L\}$ has *no influence* over A, is indicated by $\neg A$ following $\{\alpha, \beta\}||/\{K, L\}$.

Hypothesis 4 *The two agent parallel influence null hypothesis is that* $\{\alpha, \beta\}||/\{K, L\}$ *has no influence over A. Evidence for this will be:*

i) $\neg A$ *following* $\{\alpha, \beta\}||/\{K, L\}$

Examining these hypotheses against a simple two agent example using the α with a hose and β operating a pump example described above.

By inspecting figure 4.3 we observe that:

- History p_1 has $\neg A$ following $\alpha||\beta/\neg K, \neg L$. Neither agent has chosen to execute influence over A, $\neg A$ holds and, as for the single agent case, this is supporting evidence for two agent influence providing evidence that A is not constant.
- History p_2 is more difficult to classify. $\neg A$ follows the $\neg K/L$ action pair. Taken individually, A following $\neg K$ supports the single agent influence hypothesis for α and $\neg A$ following L denies the single agent influence hypothesis for β. Recall that to be looking for two agent influence the coaching agent has already been unable to show single agent influence for either α or β so this is not really

helpful information. Taken together the agents action fall outside of the K/L that the two agent parallel hypothesis requires and give $\neg A$ so this history pattern provides support.

- History p_3 has $\neg A$ following the required actions and denies the two agent parallel hypothesis.
- History p_4 is similar to p_2 and supports the hypothesis.
- History p_5 offers no evidence as do histories p_6 and p_8.
- History p_7 has the required group action leading to A and provides positive evidence.

These data are summarised in table 4.2 with the group choice indicated as $\{K, L\}$with negation of this group choice, $\neg\{K, L\}$ following definition 31. The data in table 4.2 supports our statements above.

Table 4.2: Two agent parallel influence

K	L	A	$\{K, L\}$	$\{\alpha, \beta\} \| / \{K, L\} \rightsquigarrow A$
$\neg K$	$\neg L$	$\neg A$	$\neg\{K, L\}$	Supports
$\neg K$	$\neg L$	A	$\neg\{K, L\}$	Neutral
$\neg K$	L	$\neg A$	$\neg\{K, L\}$	Supports
$\neg K$	L	A	$\neg\{K, L\}$	Neutral
K	$\neg L$	$\neg A$	$\neg\{K, L\}$	Supports
K	$\neg L$	A	$\neg\{K, L\}$	Neutral
K	L	$\neg A$	$\{K, L\}$	Denies
K	L	A	$\{K, L\}$	Supports

Parallel influence is more complex than single agent influence because – assuming a simple increase in the number of agents – of the larger set of available choices. Instead of a choice of single actions there is a choice of action combinations. Because we are considering parallel joint influence leading to A then anything where participating agents bring about A individually does not constitute joint influence.

4.7 Serial influence

Before proceeding to examine serial influence it is worth examining how we may apply STIT semantics to our discrete branching time agent framework. STIT operators may be in one of two general families, that of *deliberative* semantics and that of *achievement* semantics. Both families adopt the same underlying view of agents as choice partitions on branching time but differ in their methods of evaluating the truth of a statement. The primary conceptual difference between the two operators is that the truth of an achievement stit $[\alpha\ astit\colon A]$ depends on two separated moments whereas the deliberative STIT, $[\alpha\ dstit\colon A]$, depends only on a single moment.

An achievement STIT formula is evaluated at an outcome moment and some prior moment where the agent in question makes a choice that guarantees the outcome. Recall that the evaluation rule for the achievement STIT is: $\mathcal{M}, m/h \models [\alpha\ astit\colon A]$ iff there is a moment $w < m$ such that:

1. $\forall m' Choice^{w}_{\alpha}$ equivalent to m, $\mathcal{M}, m'/h' \models A$ for all $h' \in H_{(m')}$

2. $\exists m'' \in i_m$ such that $w < w''$ and $\mathcal{M}, m'' \not\models A$ for some $h'' \in H_{(m'')}$

A deliberative formula is evaluated at the moment where an agent makes its choice and, in the case of our system, acts and its evaluation rule is: $\mathcal{M}m/h, \models [\alpha\ dstit\colon A]$ iff

1. $\mathcal{M}, m/h, \models A \forall h' \in Choice^{m}_{\alpha}$

2. $\exists h'' \in H_{(m)}$ for which $\mathcal{M}m/h'', \not\models A$

Both of these evaluation rules share positive, clause 1, and negative, clause 2, requirements. This means that A is not settled true so that an agent's actions may be seen as having some real effect. This negative requirement is represented in our hypotheses by the requirement that there be at least one occurrence of what we term negative evidence, $\alpha/\neg K \rightsquigarrow \neg A$.

One of the difficulties of characterising behaviours involving a number of independent agents is the role of histories. If β acts in a manner which assists α then how do we treat the intersection of histories, do α and β share futures? Horty and Belnap [67] note that although the evaluation rules above refer to histories and moments this is only for semantic uniformity and the histories are idle in the achievement STIT evaluation. This bodes well for multiple agent actions as an achievement STIT may guarantee that A holds at an instant and that instant may act as the "transfer point" between agents and this is rather simpler than managing intertwined histories. The achievement approach, however, brings difficulties.

4.7.1 interaction between the two stits

Horty and Belnap explore the differences between deliberative and achievement statements by examining plausible interaction between the two types

$$[\alpha\ astit\colon A] \supset P[\alpha\ dstit\colon FA]$$

$$[\alpha\ dstit\colon FB] \supset F[\alpha\ astit\colon B]$$

4.7.2 Influence and STIT

Nested STIT expressions are a useful tool for analysing single agent behaviour. However, for various reasons nested STIT expressions present difficulties when considered in an other agent setting. Consider the expression $[\beta\ istit\colon [\alpha\ istit\colon A]]$ With β operating a valve at the tender and α operating a hose. This says that β sees to it that α sees to it that A, the equipment not burning, holds. Belnap et al. [8, Chapter 10] question how meaningful this statement is and suggest a number of interpretations. Our approach is to interpret this as a statement of influence, β influences α which has influence over A and to apply what we term an *instantaneous* STIT at the point where influence is exercised. The influence reading handles many of the objections of [8] and may be thought of as a strategic reading of such statements.

This is not the main reason for discussing STIT here, the main reason is that a nested STIT has an explicit ordering. When we say $[\beta\ istit\colon [\alpha\ istit\colon A]]$ we expect α to be the agent that finally causes A. This would not be the case if Alpha points a hose at the burning equipment and β *then* opens the foam valve. We outlined Xu's [124] discussion of the notion of a *sufficiency requirement* in section 3.3 and here we note that the nested STIT expression has an explicit ordering which clearly identifies the final actor in a chain of actions. In $[\beta\ istit\colon [\alpha\ istit\colon A]]$ the final actor is α, the agent closest to A so any descriptions that we use must be able to account for this. The simplest method is to impose an ordering on the group of agents G and the related set of agent choices C.

4.7.3 Serial influence

We begin by imposing an ordering on groups of agents and their associated groups of actions. The intuition is that for a group of agents G and an associated group of choices C we have an implicit set of moments where each agent makes its choice. Each of these moments is associated with an instant and these instants are ordered so as to provide a discrete time background for the agent world. Each agent / action pair is, thus, implicitly associated with an instant if we order these instants then we are also ordering the groups of agents and actions. We are concerned here with serial behaviour, behaviour which involves agent choices at different instants. Parallel behaviour may be viewed as a degenerate form of serial behaviour with the set of choice instants collapsing to a single instant.

Definition 33 *Given a group of agents G and a group of choices C we may generate a set of moments, M, where each agent in G makes the appropriate choice in C. Each moment in M has an associated instant I and instants have a fixed ordering. We project the ordering of these instants on to the sets off agents and choices.*

This allows us to list agents in an ordered manner. If $G = \{\alpha,\beta,\ldots,\delta\}$ then α*'s* choice precedes β*'s* choice which precedes δ and all intermediate agent choices. We wish to consider serial and parallel behaviour separately and the condition that each agent act at a unique moment prevents serial behaviours from becoming parallelised.

Having defined an ordering on agents and their choices we now discard it from our notation by assuming that ordering is implicit. This implicit ordering will allow us to use the same notation for sets of agents and sets of actions. A single agent or parallel statement will be at a single instant and a serial statement will cover a number of instants.

Definition 34 *For a temporally ordered group of agents, $G = \{\alpha, \beta,\ldots,\delta\}$ and a similarly ordered group of choices, $C = \{K \in Choice_\alpha, L \in Choice_\beta,\ldots,N \in Choice_\delta\}$ we will refer to the serial choice of C by G along a sequence of moments as $G;/C$.*

We have two ordering required is implicit in the influence expression and we continue with this assumption.

Definition 35 *We impose an implicit ordering on groups of agents and choices. The ordering used is dependent on the influence function in question.*

i) Given a group of agents G, a set of choices C and a statement that $G;/C$ has influence over A then we impose the temporal ordering of definition 33 on $G;/C$

ii) Given a group of agents G, a set of choices C and a statement that $G||/C$ has influence over A then we impose a simultaneous choice on $G||/C$ forcing all agents, $Ag \in G$ to choose at the same instant.

We now consider negation of choice in the multi agent serial case. Intuitively this negation is similar to that of the parallel case and our adoption of implicit ordering allows us to use the notion of definition 31,

Intuitively serial influence occurs when an ordered group of agents, G executing an ordered group of choices, C, has influence over A. As before we consider the negation of C as definition 31. The serial nature of influence here brings a minor complication, both G and C are ordered groups of agents and this ordering is a necessity. If the agent choices remain the same but their ordering changes then we have G' and C', these are not equivalent to G and C so any change in ordering is also considered a negation of G and C. Positive evidence, negative evidence and counter evidence are as for the parallel case. The coaching agent's two agent serial influence hypothesis is:

Hypothesis 5 *Two agent serial influence hypothesis is that $\{\alpha,\beta\}\ ;\ /\{K,L\}$ has influence over A. Evidence for this hypothesis will be:*

i) A following $\{\alpha,\beta\}\ ;\ /\{K,L\}$

ii) At least one $\neg A$ following $\{\alpha,\beta\}\ ;\ /\neg\{K,L\}$

iii) No evidence that either α or β has individual influence over A

A counter hypothesis, the two agent serial influence null hypothesis, is that $\{\alpha,\beta\}\ ;\ /\{K,L\}$ has *no influence* over A, is indicated by $\neg A$ following $\{\alpha,\beta\}\ ;\ /\{K,L\}$.

Hypothesis 6 *Two agent serial influence null hypothesis is that $\{\alpha,\beta\}\ ;\ /\{K,L\}$ has no influence over A. Evidence for this will be:*

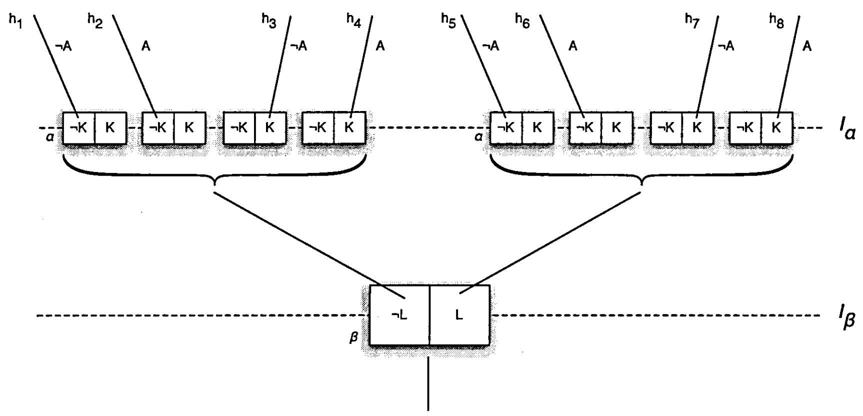


Figure 4.4: Possible histories for two agent serial influence on A

i) $\neg A$ *following* $\{\alpha,\beta\}$; $/\{K, L\}$

4.7.4 An example of serial influence

Serial influence is a further complication of parallel influence as, in a two agent case, the additional choice partitioning is augmented by ordering. In the example here we consider only one ordering, *β's* action precedes α and we are seeking to characterise [β *istit*: [α *istit*: A]] as $\{\beta, \alpha\}$; $/L, K$ has influence over A The ordering of G and C described above are explicit when these groups are written out in full, β/L precedes α/K.

An example is illustrated in figure 4.4 where β chooses from $\{\neg L, L\}$ at instant I_β and at the later instant I_α, α chooses from $\{\neg K, K\}$.

Examining each of the histories and assuming that the ordering of choices is fixed we note that:

- History h_1 supports the two agent serial influence hypothesis, the choice requirements are not met and the outcome is not satisfied and this history is a necessary requirement to indicate that in the absence of the agents that $\neg A$ is possible.

- History h_2 provides no information.

- History h_3 supports the two agent serial influence hypothesis, the behaviour requirements have not been met and $\neg A$ holds.
- History h_4 provides no information.
- History h_5 supports the hypothesis, as for h_3 the behaviour requirements have not been met and $\neg A$ holds.
- History h_6 provides no information.
- History h_7 denies the hypothesis, we observe all of the behaviour requirements being satisfied but the result is not as required.
- History h_8 supports the hypothesis and is a necessary requirement.

Representing these in table form we see that two agent serial influence hypothesis, when considered with implicit ordering, satisfies evidence requirements in the same manner as the two agent parallel influence hypothesis.

Table 4.3: Two agent serial influence

K	L	A	$\{L, K\}$	$\{\beta, \alpha\}\ ;\ /\{L, K\} \rightsquigarrow A$
$\neg K$	$\neg L$	$\neg A$	$\neg\{L, K\}$	Supports
$\neg K$	$\neg L$	A	$\neg\{L, K\}$	Neutral
$\neg K$	L	$\neg A$	$\neg\{L, K\}$	Supports
$\neg K$	L	A	$\neg\{L, K\}$	Neutral
K	$\neg L$	$\neg A$	$\neg\{L, K\}$	Supports
K	$\neg L$	A	$\neg\{L, K\}$	Neutral
K	L	$\neg A$	$\{L, K\}$	Denies
K	L	A	$\{L, K\}$	Supports

Table 4.4: Single, two agent parallel and two agent serial influence

K	L	A	$\alpha/K \rightsquigarrow A$	$\beta/L \rightsquigarrow A$	C	$G\|\|/C \rightsquigarrow A$	$G\,;/C \rightsquigarrow A$
$\neg K$	$\neg L$	$\neg A$	Supports	Supports	$\neg C$	Supports	Supports
$\neg K$	$\neg L$	A	Neutral	Neutral	$\neg C$	Neutral	Neutral
$\neg K$	L	$\neg A$	Supports	Denies	$\neg C$	Supports	Supports
$\neg K$	L	A	Neutral	Supports	$\neg C$	Neutral	Neutral
K	$\neg L$	$\neg A$	Denies	Supports	$\neg C$	Supports	Supports
K	$\neg L$	A	Supports	Neutral	$\neg C$	Neutral	Neutral
K	L	$\neg A$	Denies	Denies	C	Denies	Denies
K	L	A	Supports	Supports	C	Supports	Supports

4.8 Agent influence - discussion and generalisation

We may condense the table above by simplifying our view and consider evidence as being either complete or incomplete. Complete evidence simply means that all of the requirements of a hypothesis are met. If, for example, a parallel hypothesis requires that two agents choose K and L in order to influence A then if they do choose K and L then the evidence is complete. Any choice that is not K and L is incomplete. The

Table 4.5: Condensed agent influence

Evidence	A	Single	Parallel	Serial
Incomplete	$\neg A$	Supports	Supports	Supports
Incomplete	A	Neutral	Neutral	Neutral
Complete	$\neg A$	Denies	Denies	Denies
Complete	A	Supports	Supports	Supports

core evidence for each of the hypotheses was based on agent action. Our approach to the negation of group and sequential actions allows us to treat group and sequential actions in the same manner as single actions. From the condensed data of table 4.5 we see that this treatment of action carries from single agent cases to serial and parallel two agent cases.

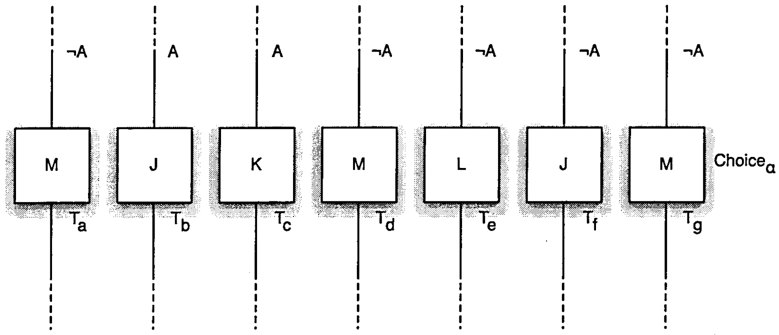


Figure 4.5: Historical data showing α's single choice at various moments

4.9 From influence hypotheses to branching time

We have represented influence in terms of hypotheses about agent influence. This is a first step and in order to be able to cast these influence notions as a reading for nested other agent STIT statements we need to be able to represent influence in branching time. In the previous section we considered the influence of a single agent, agents operating in parallel and agents operating in sequence. We follow the same sequence so as to have a parallel development of the ideas set against the different background of branching time and its more explicit representation of agent choice.

4.9.1 Representing agents

Prior [94] and Thomason's [112] branching time framework provides for an explicit representation of agent choice. Coaching agents will interpret agent history data which gives them observations of agent actions. Coach observations will be of a single agent / choice pair leading to a single history for that agent. A coaching agent may, for a single agent α observe a series of choices, $T_a \ldots T_g$ as illustrated in figure 4.5 representing several events which a coach will have observed at different times and have been aggregated. These events show what choices are available to α and data associated with α's postcepts following the choice. Note that here we have limited the postcept data associated with each history to A or $\neg A$ as this is the notional vehicle for describing α's influence. A is simply a subset of the agent's postcepts. These data

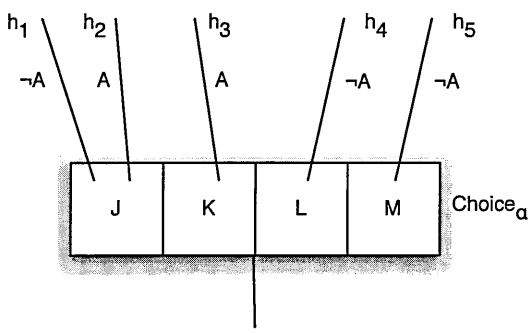


Figure 4.6: Aggregated historical data represents α's choice set and observed outcomes

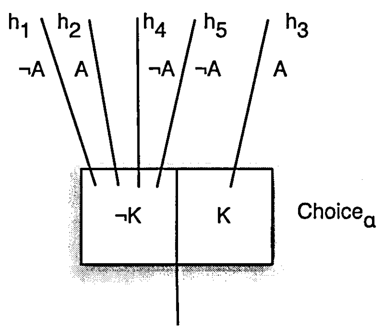


Figure 4.7: Binary representation of α's choices relative to K

may be aggregated by a coaching agent so as to represent an agent's potential choices and possible results related to those choices. Following aggregation the coach will have a set of choices which it has observed α making and this set of choices is treated as an abstract agent class. The histories of figure 4.5 may be aggregated to give the agent image of figure 4.6 where we consider α as a set of choices, $Choice_\alpha$. Here $Choice_\alpha = \{J, K, L, M\}$ leads to five possible histories. If we apply the same notion of negation as definition 27 we see this collapse into two choices as illustrated in figure 4.7. All of the choices that are $\neg K$, that is $\{J, L, M\}$ are grouped in a single $\neg K$ equivalence class leaving K and any histories associated with it in a complementary equivalence class. However, our interest here is focused on evaluating α's ability to bring

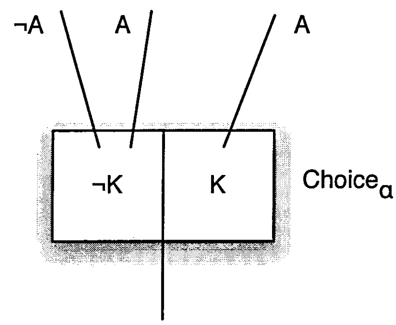


Figure 4.8: Condensed binary representation of *α's* choices relative to *K*

about A and the four histories branching forwards from the condensed $\neg K$ choice partition are redundant. What is of interest is that the $\neg K$ choice leads to possible futures where A holds and where A does not hold. We condense the representation further to remove redundancy and give an outline of the minimal agent class than a coaching agent will use for behaviour analysis and synthesis in figure 4.8.

As a brief aside, consider the computational load on coaching agents. The large branching factor evident in figure 4.6 brings formidable computing requirements. The focus on A with the consequent reduction in branching factor concomitant with the binary choice may significantly reduce computational complexity. Coaching agents will need to maintain and update a database of full agent class representations, as in figure 4.6, but any analysis and behaviour synthesis will be based on a condensed binary choice representation of an agent. We define an observed agent and condensed binary agent here:

Definition 36 *Given a set of agent histories H each containing a number of choices* $c \in H$ *then we may construct an observed agent class and represent it as a set of choices* $C = \{c : c \in H\}$.

We define the condensed binary choice agent where $\alpha / K \rightsquigarrow A$.

Definition 37 *Given an abstract agent class* α *with a set of choices C and a single choice element* $K \in C$ *we define a binary abstract agent for* α / K *as an agent represented by the choice partition* $\{C \backslash K, K\}$ *and write this as* $\{\neg K, K\}$.

This representation of agents as a choice partition with its association with histories provides a method for incorporating the agent $\alpha/K \rightsquigarrow A$ concept discussed above into a branching time framework. A framework that provides a link to the STIT representation that we intend develop. Before addressing our evidence hypotheses in this context we briefly mention some related issues which will be discussed in more detail later.

4.10 A practical theory of influence?

In the previous chapter we discussed the notion of agent influence in general terms noting that it may provide a way to make sense of nested other agent statements without additional system or society burdens. In this chapter we moved this theoretical consideration onto a more practical grounding, one based on observation. This led us to describing agents in terms of equivalence classes based on their observed choices. From these agent and choice pairs we are able to detect influence by its effect on the results or availability of agent choice. We can also build hypotheses, based on observed behaviour, about agent ability and we are able to classify observations into one of three classes which support, deny or are neutral towards a hypothesis. Practicality, here, means that from observations we are able to hypothesise about ability. Our next step is to examine agent influence in a formal setting. We have frequently compared influence to STIT and to ground this comparison we present a partial logical characterisation of influence in the following chapter. After characterising influence we return to a practical setting to discuss the operation of coaching agents, already mentioned frequently, and build a simple system to test the basic ideas that we have presented.

Chapter 5

A partial logical characterisation of influence

We have explored the notion of hypotheses of agent influence based on observed agent behaviour and have identified influence in contexts of serial and parallel behaviour. We have noted that influence is not as strong as STIT but that there are similarities and that influence tends to STIT with strong evidence. One of the results of STIT being widely researched is that it has been has been characterised logically and this provides a sound basis for reasoning with and manipulating STIT expressions. In preparation for a discussion of how coaching agents will operate, essentially how they may manipulate hypotheses, we consider how a number of standard modal rules apply to evidence based hypotheses. We also introduce an *other agent extension* to certain of the standard rules so as to explicitly investigate the workings of other agent nested statements.

Our contention is that our influence operators are similar to STIT and that given perfect evidence they are the same, essentially that STIT is a subset of influence. Before exploring this claim we outline some supporting concepts to provide us with relation operators and a means of ranking evidence by its strength. We discuss our evidence based relations in a little more detail before outlining our *other agent* extension and considering a number of modal rules.

5.1 Preliminaries, evidence based relations

We have a notion of agent ability for single agent cases and for group cases where agents operate in serial or parallel. If an agent, α has self contained ability then we say that an appropriate action by α will lead to A holding and we outlined a notation for this in definition 29. If, however, α*'s* ability is contingent on another agent's action, an action which may be before or at the same time as α*'s* action, then we say that α*'s* action *may* lead to A holding and building on definition 29 we introduce a *may lead to* operator:

Definition 38 *Given an agent, α with a choice K and a proposition A if we observe instances of $\alpha/K \leadsto A$ and instances of $\alpha/K \not\leadsto A$ we infer that α/K may lead to A and write this as $\alpha/K \Diamond\!\!\leadsto A$.*

Our may lead to relation combines the modal possibility operator, $\Diamond$, and a leads to relation, $\leadsto$. $\alpha/K \Diamond\!\!\leadsto A$ reads α/K may lead to A and indicates that α*'s* ability may be contingent on other agent actions.

In section 4.4 we observed that evidence may be divided into three classes, positive evidence (which has two requirements), counter evidence and neutral evidence. What do we consider as strong evidence and weak evidence? In an ideal, noise free environment with perfect observation absolute agent influence would always cause positive evidence. In a noisy or uncertain environment we may expect to observe some counter evidence to that agent's ability and does not necessarily mean that the agent has no influence. Maybe $\alpha's$ ability is contingent on another agent or maybe another agent acted so as to counter $\alpha's$ choice. Evidence is countable, an instance of A following α/K would constitute one piece of evidence for the hypothesis that $\alpha/K \leadsto A$. Similarly an instance of $\neg K$ following α/K would constitute one piece of counter evidence for the hypothesis that $\alpha/K \leadsto A$.

We adopt two metrics which allow us to rank hypotheses. The first of these is the ratio between the positive and negative evidence counts, the $P\!:C$ ratio which we term the *PCR*. This is a strong indicator of influence but experiments showed that in very noisy environments noise could swamp the $P\!:C$ ratio for nested behaviours. We adopt a second metric which is simply the raw difference between the positive and counter evidence tallies for a hypothesis. This $P-C$ or *PMC* value serves as a filter for noisy *PCR* data. Intuitively hypotheses for strong influence will rise to the top of a hypothesis table ranked by $P\!:C$ ratio.

Hypotheses with no influence will sink to the bottom and the *PCR* is used to identify potential candidates for other agent behaviour from the mid regions of a *PCR* ranked table. These concepts are discussed in more detail in chapter 6 but are introduced here as the notions of evidence strength in terms of $P:C$ ratio are used in the following chapter.

Note that in the text below we simply use P/C to generate the $P:C$ ratio, this allows it to reach infinity which is clearer and more intuitive for reading. In experimental code we use $P/(P+C)$ so as to limit the upper value to 1 and prevent divide by zero errors.

5.1.1 Two influence operators

In section 5.1 we introduced two relations which represent an agent being able or possibly being able to bring a proposition about. The second of these, a *may lead to* relation, could, we noted, be an indicator of other agent involvement in an agent's influence. In this section we examine these relations in a modal context with the investigation being driven by *observed evidence* so as to represent what situated coaching agents will perceive. The first, $\leadsto$ a leads to operator, may be considered as being similar to a STIT operator. For $\leadsto$ to hold we must observe evidence of ability and consistent results. This evidence of ability has a negative condition allowing us to cast this in the deliberative STIT mould. The leads to relation differs from STIT in that it is based entirely on observation and does not necessarily represent a complete characterisation of an agent's choices or exploration of its world.

The second of these, $\Diamond\!\!\leadsto$ a may lead to operator, was introduced to allow us to hypothesise about other agent influence. Casting this as a STIT operator is more awkward, $\Diamond\!\!\leadsto$ implies ability without consistency. It has positive and negative requirements just like $\leadsto$ and deliberative STIT operators but it also admits cases where an agent fails to bring something about. The intuition is simply that if there is another influence at play then the coach will observe inconsistent results. The $\leadsto$ and $\Diamond\!\!\leadsto$ operators are related, for $\leadsto$ to hold there must be no counter evidence, no cases where α on its own fails to bring about A. If there is counter evidence, a case where α does not bring about A and there is evidence of potential influence then this allows the coach to use the $\Diamond\!\!\leadsto$ operator. In a noisy and dynamic environment a coach may only rarely observe a $\rightsquigarrow$ hypothesis being satisfied. There may be false positives because of the multi agent nature of the environment. It may be that if α and β are in the same location and β acts so as to bring about A. α will see A in its set of postcepts and it may appear that α's action brought about A when it really had no effect.

Of course there are cases where an agent may not have influence. A may come about coincidentally with an α's actions. A coaching agent may then hypothesise that another agent is doing something that allows α to see to it that A holds whereas $\alpha's$ choices may be driven by the same cues that trigger A. If the coach synthesises a new behaviour based on this hypothesis and seeds it in the environment there may be more occurrences of the hypothesised other-agent action but, in this case, no change in the strength evidence for $\alpha's$ influence and the hypothesis may become a candidate for being withdrawn.

5.1.2 Modal rules – introduction

Observation is rarely perfect and environments may be noisy. It may be that α has influence over A and that with noise and uncertainty removed from evidence then $\rightsquigarrow$ holds. This will occur only in ideal circumstances. Coaching agents will admit the uncertainty in evidence by considering that a behaviour with a $P:C$ ratio above a certain level is considered as satisfying $\rightsquigarrow$. A $P:C$ ratio of ∞ indicating that the requirements of $\rightsquigarrow$ have been fully satisfied.

In order for these operators to be manipulated as modal expressions we must demonstrate that they satisfy a number of rules which are associated with characteristics of modal systems. One of the major aims of this work is to characterise other-agent influence and to do this we extend some standard modal axioms, those where it might be said that an agent influences itself, to encompass other-agent influence.

Since our system is driven by evidence and evidence is unlikely to be noiseless we approach this from two points of view. We assume, first, the $\rightsquigarrow$ operator holds then consider examples or counterexamples on this assumption. Evidence, in this case, is noiseless and that where an agent has influence over something then this influence is unambiguously captured. These ideal cases serve to demonstrate that the system is capable of supporting or denying certain modal principles. We then consider the same examples with noisy data using the $\Diamond\!\leadsto$ operator to investigate if it is safe to consider whether or not the same modal principles are supported by an influence operator.

We have observed that with a $P\!:C$ ratio of ∞ is indicative of a $\leadsto$ type expression. In a noisy data environment this is unlikely to occur and this, potentially, brings difficulties. To overcome – though it is perhaps more accurate to say accommodate – these difficulties we introduce two hurdle values for the $P\!: C$ ratio. A high hurdle above which we consider $\leadsto$ to be the case, a low hurdle below which we hold that the agent has no influence and between the hurdles where we consider $\Diamond\!\leadsto$ to be the case.

Standard STIT semantics are choice agnostic, they read that an agent is able to see to it that A holds in a non prescriptive manner. The $\leadsto$ and $\Diamond\!\leadsto$ operators are similar but we bring an implicit prescriptive element their association with hypotheses based on the observation of agent choice. This is a necessary requirement which allows coaching agents to meaningfully analyse and manipulate agent behaviour. It may well be that an agent has two different actions that lead to the same result, this means that coaching agents may hold two hypotheses – one for each action. When we consider hypotheses in conjunction with a $\leadsto$ or $\Diamond\!\leadsto$ relation then the agent and action are implicit parts of the statement.

5.1.3 Syntax for a leadsto operator

We have discussed leadsto and STIT separately, we now bring these together and cast our *leads to* and *may lead to* operators as STIT like operators. We do this by replacing the generic STIT part of a STIT expression with either $\leadsto$ or $\Diamond\!\leadsto$ and read the expressions as follows:

Definition 39 *Substituting a $\leadsto$ expression for* STIT *in a* STIT *expression yields statements of the form* $[\alpha \leadsto\!: A]$ *and* $[\alpha\ \Diamond\!\leadsto\!: A]$ *which are read, respectively, as α has a choice which will lead to A and α has a choice which may lead to A.*

The readings suggested above give no indication of which STIT family that the *leads to* operators belong to and give no indication as to what the agent's choice really is. At this point we address the STIT family question by taking a pragmatic approach and this is that agents have a choice which will or may lead to A

at some instant following the agent's choice. The actuality of an agent's choice is resolved by one or more related coach hypotheses.

The rules discussed below help to describe expectations and requirements of system behaviour and the properties of our $\rightsquigarrow$ and $\Diamond\!\!\rightsquigarrow$ operators. These rules allow us to characterise the system and operators in a modal framework.

5.1.4 The other-agent extension

In the standard, single agent, versions of many of the rules considered in this characterisation there are a number of sentences which may admit multiple agents. The standard modal axiom C, for example, states that, and using $\rightsquigarrow$ in place of STIT, $[\alpha \rightsquigarrow: A] \wedge [\alpha \rightsquigarrow: B] \supset [\alpha \rightsquigarrow: A \wedge B]$ and we consider multi agent extensions of this. The standard single agent / multiple sentence statement becomes a multiple agent / single sentence statement. Casting C in this mould replacing the sentences, A and B with agents α and β and introducing a single sentence gives C_{agent}. C_{agent}, then, states that $[\alpha \rightsquigarrow: A] \wedge [\beta \rightsquigarrow: A] \supset [\alpha \| \beta \rightsquigarrow: A]$. These agent extensions need to be considered in the parallel case, as above, and the serial case. C_{agent}, in serial form states that $[\alpha \rightsquigarrow: A] \wedge [\beta \rightsquigarrow: A] \supset [\alpha;\beta \rightsquigarrow: A]$. Jumping ahead briefly, where C_{agent} holds it appears to indicate that α and β's actions are not mutually exclusive. Such agent extensions of standard rules pose a problem, it is unreasonable to expect that a system should validate a rule like C_{agent} without context and where required we provide illustrative examples.

This completes the preliminary discussion and we now consider a number of modal rules.

5.2 Influence with modal rules and axioms

We listed some standard modal rules and axioms in section 2.6.6 and in this partial characterisation we examine a number of these rules and axioms with $\rightsquigarrow$ and $\Diamond\!\!\rightsquigarrow$ operators substituted for generic STIT operators. This is intended to give an outline of how closely our influence operators follow standard STIT. The other agent extension is used where appropriate and we consider whether or not observations will allow us to carry standard rules into the domain of other agent actions.

5.2.1 Negative necessitation, $\overline{N}$

We first consider the negative necessitation rule. The import of this is that an agent should not be able to influence $\top$. Since $\top$ is a constant we should expect the rule in equation 5.1 to be valid.

$$\overline{N}.\ \neg[\alpha \rightsquigarrow: \top] \tag{5.1}$$

We assume that a counterexample holds, N. $[\alpha \rightsquigarrow: \top]$ and that α has some choice K available to it that has influence over $\top$. For this to be the case we would need to observe positive evidence where $\alpha/K \rightsquigarrow \top$ and our evidence requirements demand that we observe at least one instance of $\alpha/\neg K \rightsquigarrow \bot$. However, since $\top$ is a constant we will never see $\alpha/\neg K \rightsquigarrow \bot$. A coaching agent collecting observations may see data like that of table 5.1 where the $\times$ denoting the lack of negative influence indicates that α/K has no influence over $\top$.

Table 5.1: Evidence for $\alpha/K \rightsquigarrow \top$

Hypothesis	Evidence			P:C ratio	Conclusion
	Positive-	Negative-	Counter-		
$\alpha/K \rightsquigarrow \top$	n	$\times$	0	0	No influence

The impossibility of generating a counterexample indicates that $\overline{N}$ is valid. Note, also, that because of the impossibility of generating a counterexample, a necessary requirement for elevating a hypothesis, it is neither necessary nor possible to move to the $\Diamond\!\!\rightsquigarrow$ operator.

5.2.2 Rule of necessitation, *RN*

For coaching agents to operate meaningfully the system must support some form of inference. This is represented in modal logic by *RN*, the rule of necessitation. *RN* and the distribution axiom, known as *K* after Saul Kripke, together are widely accepted as being the most basic properties of a modal system (see

Chellas [30, page 6]). Chellas [30, page 245] represents *RN* as shown in equation 5.2.

$$RN.\ \frac{A}{\Box A} \tag{5.2}$$

Casting this in our influence framework gives equation 5.3.

$$RN.\ \frac{[\alpha \rightsquigarrow : A]}{A} \tag{5.3}$$

RN serves as an indicator of individual ability and may be interpreted as a trigger for what we term *elevating* a hypothesis. Recalling the notion of influence domains and gateways, illustrated in figure 3.2, we make the following observation:

Observation 15 *Suppose a hypothesis that an agent, α, has influence over a proposition, A. If the observed evidence supports the hypothesis but also has a significant counter evidence component, then a coaching agent may generate further hypotheses that other agents play a part in the cases where $[\alpha \rightsquigarrow : A]$. The coaching agent will "elevate" this hypothesis into the next behaviour domain by generating other agent hypotheses which support its observations.*

If an agent has a complete individual influence to bring about A, let us call this an innate influence, then following that agent's making the appropriate choice then A will hold and we would observe positive and negative evidence and no counter evidence. If an agent lacks innate influence but is able to bring about A either in conjunction with another agent or following some action by another agent then we will observe positive and negative evidence with some counter evidence. If an agent has no influence then we would observe no positive evidence, some negative evidence and some counter evidence. These scenarios are for noiseless single agent environments. Staying in a noiseless, single agent environment then a counterexample to $[\alpha \rightsquigarrow : A] \supset A$ would be the second example above, where an agent's influence is contingent. Possible observed evidence for *RN* is illustrated in table 5.2.

Counterexamples to *RN* with the $\rightsquigarrow$ operator fall into two classes, one which satisfies the $\Diamond\!\!\rightsquigarrow$ operator and allows a coaching agent to elevate a hypothesis and one which falsifies both the $\rightsquigarrow$ and $\Diamond\!\!\rightsquigarrow$ operators

Table 5.2: Evidence and *RN*

Hypothesis	Evidence			*P*:*C* ratio	Conclusion
	Positive-	Negative-	Counter-		
(5.2a): $\rightsquigarrow$ example					
$\alpha/K \rightsquigarrow A$	n	✓	0	∞	$\alpha/K \rightsquigarrow A$
(5.2b): $\Diamond\!\!\rightsquigarrow$ example, counterexample to $\rightsquigarrow$					
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K \Diamond\!\!\rightsquigarrow A$

allowing a coach to discard a hypothesis. Falsification of the $\Diamond\!\!\rightsquigarrow$ operator occurs when the counter evidence tally is significantly greater than the positive evidence tally. This level will be heavily dependent on the agent's context and we suggest that this be addressed by ranking hypotheses and discarding the weakest in the ranked list.

RN is sufficiently strong for us to build on it and go from necessitation to inference placing us in a position where we may begin considering other agent constructs.

5.2.3 Rule of inference *RR*

RR provides a simple rule which allows coaching agents to infer results from sequences or combinations of observations across two agents or actions. This is sufficient for coaching operation described in chapter 6 and the the experiments of chapter 7. Here we consider two propositions as a conjunction. *RR* is a general rule and we apply an other agent extension to refine *RR* to deal specifically with serial or parallel other agent choices.

$$RR. \ \frac{(A \wedge B) \rightarrow C}{(\Box A \wedge \Box B) \rightarrow \Box C} \tag{5.4}$$

Our interpretation or *RN* carries a two implicit requirements, the first of these is the assumption that a specific agent choice leads to A holding. This assumption of choice is based on our notion of condensed binary choice, definition 37 so that all "other choices" are grouped together. The second requirement is that the propositions under consideration are not constant. This is one of the foundations of our theory of influence and means that we will have observed at least one instance of the negation of the proposition under consideration following one of the agent's other choices. Returning, briefly, to *RN*, when we write $\frac{[\alpha \rightsquigarrow : A]}{A}$ we are saying that some choice byα leads to A being true and that given that choice by α we may infer that A holds.

Bringing the standard *RR* of equation 5.4 into our influence setting we rewrite it as equation 5.5.

$$RR. \quad \frac{([\alpha \rightsquigarrow : A] \wedge [\alpha \rightsquigarrow : B]) \rightarrow [\alpha \rightsquigarrow : C]}{(A \wedge B) \rightarrow C} \tag{5.5}$$

This states that if *α's* influence brings about A and *α's* influence brings about B implies that *α's* influence brings about C then, given the assumptions immediately above, we may infer that $A \wedge B$ implies C.

Our concern is to show that *RR* holds for three cases where $(A \wedge B) \rightarrow C$, that of a single agent, that of two agents acting in series and that of two agents acting in parallel. Before considering *RR* we make some assumptions about the environment. These are that:

(i) A and B are independent propositions.

(ii) C occurs spontaneously when A and B hold.

(iii) For C to occur then either A and B must be brought about at the same instant or A must already hold when B is brought about. This represents parallel and serial action.

(iv) Agents may execute only one choice at an instant.

(v) A and B are persistent with lifespan of at least one cycle.

This list of assumptions provides for a limited set of circumstances for testing *RR* and we, later, briefly discuss a more general setting. Item (iii) imposes an ordering requirement for the serial case, this is simply as an illustrative convenience. The variable lifespan for A and B, item (v) introduces an element of noise. Note that in the noiseless case this lifespan is for as long as a proposition is required. Item (iv) means that for a single agent the conjunction of actions must be serial and (v) allows for a single agent, α, to bring about A.

We restate equation 5.5 to reflect the implicit single agent ordering requirement of item (iii):

$$RR.\ \frac{([\alpha \rightsquigarrow: A]; [\alpha \rightsquigarrow: B]) \rightarrow C}{(A; B) \rightarrow [\alpha \rightsquigarrow: C]} \tag{5.6}$$

Equation 5.6 states that in order to be able to bring about A then α must already have brought about A before bringing about B and this ordering follows the requirements discussed in section 4.7. In line with the implicitly prescriptive nature of hypotheses we associate an agent choice with each proposition giving two hypotheses; $\alpha/K \rightsquigarrow A$ and $\alpha/L \rightsquigarrow B$

Table 5.3: Evidence and single agent *RR* in a noise free setting

Hypothesis	Evidence			*P*:*C* ratio	Conclusion
	Positive-	Negative-	Counter-		
$\alpha/K \rightsquigarrow A$	n	✓	0	∞	$\alpha/K \rightsquigarrow A$
$\alpha/L \rightsquigarrow B$	p	✓	0	∞	$\alpha/L \rightsquigarrow B$
$\alpha/K \rightsquigarrow C$	q	✓	r	q/r	$\alpha/K \Diamond\!\!\rightsquigarrow C$
$\alpha/L \rightsquigarrow C$	s	✓	t	s/t	$\alpha/L \Diamond\!\!\rightsquigarrow C$
$\{\alpha/K; \alpha/L\} \rightsquigarrow C$	u	✓	0	∞	$\{\alpha/K; \alpha/L\} \rightsquigarrow C$
$\{\alpha/K; \alpha/L\} \rightsquigarrow A$	u	✓	0	∞	$\{\alpha/K; \alpha/L\} \rightsquigarrow A$
$\{\alpha/K; \alpha/L\} \rightsquigarrow B$	u	✓	0	∞	$\{\alpha/K; \alpha/L\} \rightsquigarrow B$
$\{\alpha/L; \alpha/K\} \rightsquigarrow C$	v	✓	w	v/w	$\{\alpha/L; \alpha/K\} \Diamond\!\!\rightsquigarrow C$

In the noiseless case of table 5.3 we see that there are n instances of α/K bringing about A and p instances of α/L bringing about B. Because the environment is noiseless there is no counter evidence for these cases and the single agent hypotheses are satisfied. We may also observe cases of α/K and α/L bringing about C. The latter is necessary because that is the action that, in conjunction with A holding, brings about C. This is a noise free environment but C may hold before α/K so instances of C following α/K will be observed. In both of these cases there will be counter evidence indicating that the single agent hypothesis does not satisfy the requirements for bringing about C and causing the coach to elevate these hypotheses by seeking other agent influence. The other agent, in this case, is the same α but there is a requirement that it sequence its choices. Considering A, B and C as subjects of serial hypotheses on $\{\alpha/K;\alpha/L\}$ we may observe u instances (with $u <= n$ and $u <= p$) of serial action bringing about A, B and C. The evidence levels (in this noise free environment) will be the same for each since (by items (ii) and (iii)) which gives three hypotheses with different results and the same evidence levels. We can immediately discard the serial action hypotheses for A and B as we already have satisfied single agent hypotheses for these propositions and there is no point in seeking complex explanations for cases that have been dealt with, this leaves us with a single hypothesis which satisfies equation 5.6 in a noise free setting. The out of order hypothesis, $\{\alpha/L;\alpha/K\} \leadsto C$, will present some positive evidence but negative evidence will be greater leading to a *may lead to* conclusion with weak evidence.

RN allows for the elevation of a hypothesis into a two agent domain and *RR* allows this extension so that a coaching agent may generate two agent hypotheses.

In a noisy setting, one where other agents may have actions that counter $\alpha's$ abilities, we would expect to see counter evidence in each case. The differentiation between hypotheses must then be on evidence strength rather than the difference between perfect and imperfect evidence and possible observations are listed in table 5.4 which starts from the same *leads to* hypotheses as table 5.3. Here we would see the $P:C$

Table 5.4: Evidence and single agent *RR* in a noisy setting

Hypothesis	Evidence			P:C ratio	Conclusion
	Positive-	Negative-	Counter-		
$\alpha/K \leadsto A$	n	✓	p	n/p	$\alpha/K \Diamond\!\!\leadsto A$
$\alpha/L \leadsto B$	q	✓	r	q/r	$\alpha/L \Diamond\!\!\leadsto B$
$\alpha/K \leadsto C$	s	✓	t	s/t	$\alpha/K \Diamond\!\!\leadsto C$
$\alpha/L \leadsto C$	u	✓	v	u/v	$\alpha/L \Diamond\!\!\leadsto C$
$\{\alpha/K;\alpha/L\} \leadsto C$	w	✓	x	w/x	$\{\alpha/K;\alpha/L\} \Diamond\!\!\leadsto C$
$\{\alpha/K;\alpha/L\} \leadsto A$	y	✓	z	y/z	$\{\alpha/K;\alpha/L\} \Diamond\!\!\leadsto A$
$\{\alpha/K;\alpha/L\} \leadsto B$	a	✓	b	a/b	$\{\alpha/K;\alpha/L\} \Diamond\!\!\leadsto B$
$\{\alpha/L;\alpha/K\} \leadsto C$	c	✓	d	c/d	$\{\alpha/L;\alpha/K\} \Diamond\!\!\leadsto C$

ratios for $\alpha/K \leadsto A$ and $\alpha/L \leadsto B$ being relatively good compared with those of $\alpha/K \leadsto C$ and $\alpha/L \leadsto C$. Once again we may discard $\{\alpha/K;\alpha/L\} \leadsto A$ and $\{\alpha/K;\alpha/L\} \leadsto B$. The weaker hypotheses indicate that there is some influence and the relative strength will cause the coaching agent to elevate the hypotheses to a set of other agent hypotheses.

Bringing in the other agent extension we may replace one of the instances of α with another agent. This other agent may be another instance of an α class agent or another agent that has an appropriate equivalent choice. The independence of A and B and implicit sequencing of actions in the single agent case make the serial other agent version of RR a simple substitution. Item (iii) in our list of conditions states that simultaneous action by two different agents will bring about C. This may be thought of a statement of the physics of this agent world and if we did not have this condition then this simultaneous action would not lead to C in the agent world and there would be no evidence or extremely weak evidence supporting such a hypothesis.

5.2.4 The convergence axiom, C

Chellas [30, page 20] lists this principle as:

$$(\Box A \wedge \Box B) \rightarrow \Box(A \wedge B) \tag{5.7}$$

If an agent were to see to it that A holds and that B holds at an instant and its choice of action is the coincidental result of its having two independently reasoned goals then it does so without intending to see to it that A and B hold jointly. Neglecting the agent's intent, however, it would be difficult to deny that the agent does see to it that A and B do hold jointly and that the principle stated in equation 5.8 is supported. A noteworthy point here is that our operators are constrained to an agent and choice pair, here we say that $\alpha/K \rightsquigarrow A$ and $\alpha/K \rightsquigarrow B$ and the implication is that it is the same choice, K, that brings about both A and B. We consider that equation 5.8 holds and attempt to generate counterexamples.

$$C.\ \ [\alpha \rightsquigarrow : A] \wedge [\alpha \rightsquigarrow : B] \rightarrow [\alpha \rightsquigarrow : A \wedge B] \tag{5.8}$$

Table 5.5: Evidence supporting C

Hypothesis	Evidence			P:C ratio	Conclusion
	Positive-	Negative-	Counter-		
(5.5a): $\rightsquigarrow$ noiseless setting					
$\alpha/K \rightsquigarrow A$	n	✓	0	∞	$\alpha/K \rightsquigarrow A$
$\alpha/K \rightsquigarrow B$	n	✓	0	∞	$\alpha/K \rightsquigarrow B$
$\alpha/K \rightsquigarrow A \wedge B$	n	✓	0	∞	$\alpha/K \rightsquigarrow A \wedge B$
(5.5b): $\Diamond\!\!\rightsquigarrow$ noisy setting					
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K \Diamond\!\!\rightsquigarrow A$
$\alpha/K \rightsquigarrow B$	q	✓	r	q/r	$\alpha/K \Diamond\!\!\rightsquigarrow B$
$\alpha/K \rightsquigarrow A \wedge B$	s	✓	t	s/t	$\alpha/K \Diamond\!\!\rightsquigarrow A \wedge B$

Considering the noiseless case and assume that $\alpha/K \rightsquigarrow A$ and $\alpha/K \rightsquigarrow B$ both hold and considering them separately we may see evidence like that presented in table 5.5a. Because the result of α/K is A and B regardless of the agent's intention then the positive evidence count will be the same in each case. We introduce another hypothesis, a hypothesis that $\alpha/K \rightsquigarrow A \wedge B$. For this hypothesis to fail then there must be at least one instance of counter evidence where either $\alpha/K \rightsquigarrow \neg A \wedge B$. or $\alpha/K \rightsquigarrow A \wedge \neg B$. This would mean that the positive evidence count for $\alpha/K \rightsquigarrow A \wedge B$. would differ from that of $\alpha/K \rightsquigarrow A$ and $\alpha/K \rightsquigarrow B$ which is not possible given the behaviour of α/K described above.

The noisy case, illustrated in table 5.5b, is not so clear. We do not have the lack of influence over B that will allow us to generate a counterexample to falsify C. Because of the noise we are unable to guarantee that evidence tallies for $\alpha/K \rightsquigarrow A$ and $\alpha/K \rightsquigarrow B$ will be the same though we may expect them to be of the same magnitude.

5.2.5 C_{agent}

This is an other agent extension, as outlined in section 5.1.4, to the standard C axiom represents scenarios where a number of agents with similar abilities exercise their abilities either in parallel or in series. This gives two versions, equation 5.9 says that if α can see to it that A and if β can see to it that A then α and β

acting simultaneously can see to it that A holds. Equation 5.10 says that if α can see to it that A and if β can see to it that A then α and β acting serially can see to it that A holds.

$$C_{\|agent}.\ \ [\alpha \rightsquigarrow: A] \| [\beta \rightsquigarrow: A] \rightarrow [\alpha \| \beta \rightsquigarrow: A] \tag{5.9}$$

$$C_{;agent}.\ \ [\alpha \rightsquigarrow: A] \wedge [\beta \rightsquigarrow: A] \rightarrow [\alpha ; \beta \rightsquigarrow: A] \tag{5.10}$$

$C_{;agent}$ is a potentially dangerous property. Consider agent actions which are mutually exclusive, either α or β may bring about A. We consider the parallel action case first, if α and β simultaneously act so as to bring about A then A will hold but it will do so as a result of *α's* action or *β's* action and not as a result of both actions. We assume that both agents have different abilities, that is that they execute different choices but these choices are functionally equivalent as far as A is concerned. If, say, $\alpha/K \rightsquigarrow A$ and $\beta/L \rightsquigarrow A$ and $\neg A$ holds then when agents act simultaneously both will perceive that A holds after their action. Inspecting the

Table 5.6: Evidence and parallel C_{agent}

Hypothesis	Evidence			P:C ratio	Conclusion
	Positive-	Negative-	Counter-		
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K \Diamond\!\!\rightsquigarrow A$
$\beta/L \rightsquigarrow A$	q	✓	r	q/r	$\beta/L \Diamond\!\!\rightsquigarrow A$
$\{\alpha/K \| \beta/L\} \rightsquigarrow A$	s	✓	t	s/t	$\{\alpha/K \| \beta/L\} \Diamond\!\!\rightsquigarrow A$

data of table 5.6 we see evidence for C_{agent}. Even in a single cell world where agents are always collocated the number of instances of $\alpha/K \rightsquigarrow A$ and $\beta/L \rightsquigarrow A$ will be greater than those of $\{\alpha/K \| \beta/L\} \rightsquigarrow A$. Assuming an even distribution of noise the $P:C$ ratios n/p and q/r will be greater then s/t and ranking these as per the discussion in section 5.1 makes the single agent hypotheses appear to be more influential than the parallel action hypothesis. Returning to agents, we have two hypotheses for each – a single agent hypothesis that indicates an ability to bring about A and an other agent hypothesis indicating the same. We discussed the notion of gateways between domains of influence in section 3.2 and illustrate the single and two agent case in figure 5.1. The single agent action is contained in the single agent influence domain which is, by extension,

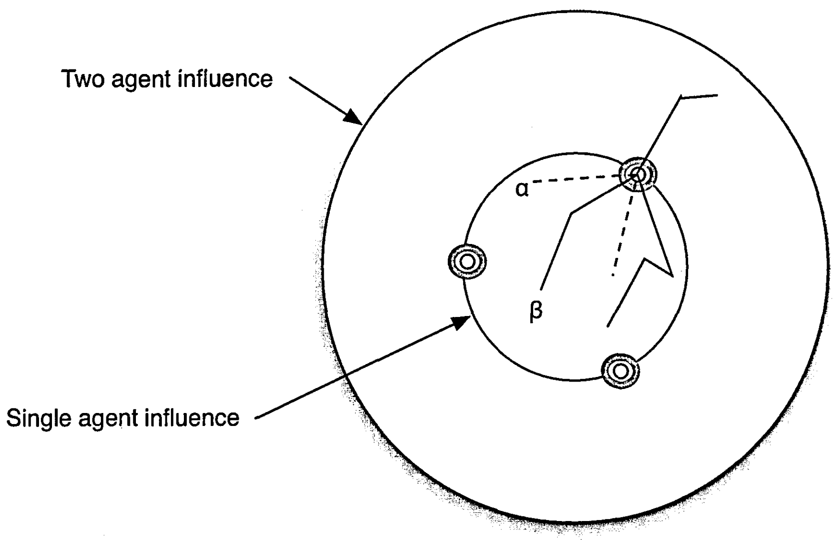


Figure 5.1: Single agent and other agent hypothesis domains

contained in the two agent influence domain. After ranking the simpler single agent hypotheses are seen to carry more influence and offer a better account of α and β's individual ability to influence A than the two agent hypothesis.

If α and β operate sequentially then coaching agents will see much more evidence of influence for single agent action than of serial action. Given $\alpha/K \leadsto A$ immediately followed by $\beta/L \leadsto A$ the latter action will, except when noise intervenes, show no influence as A already holds and no change will be evident following β/L.

Whilst we do not falsify C_{agent} we see that the foundations of our influence theory - observations of agent behaviour – do not lead to circumstances where C_{agent} may be considered as valid.

5.2.6 Whatever is necessary is the case, M

Chellas [30, page 20] lists this principle as:

$$M.\ \Box(A \wedge B) \rightarrow (\Box A \wedge \Box B) \tag{5.11}$$

M claims that if an agent is responsible for seeing to $A \wedge B$ then it is responsible for seeing to A individually and seeing to B individually.

$$M.\ [\alpha \rightsquigarrow : A \wedge B] \rightarrow ([\alpha \rightsquigarrow : A] \wedge [\alpha \rightsquigarrow : B]) \tag{5.12}$$

This is an objectionable claim and as a counter model we show that it is possible for the left hand side to hold when one of the two expressions on the right hand side is falsified. We assume that α/K has influence over A and has no influence over B which remains true.

Table 5.7: Counterexample to M

Hypothesis	Evidence			$P{:}C$ ratio	Conclusion
	Positive-	Negative-	Counter-		
(5.7a): $\rightsquigarrow$ counterexample, noiseless					
$\alpha/K \rightsquigarrow A$	n	✓	0	∞	$\alpha/K \rightsquigarrow A$
$\alpha/K \rightsquigarrow B$	n	✕	0	∞	No influence
$\alpha/K \rightsquigarrow A \wedge B$	n	✓	0	∞	$\alpha/K \rightsquigarrow A \wedge B$
(5.7b): $\Diamond\!\!\rightsquigarrow$ counterexample, noisy					
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K \Diamond\!\!\rightsquigarrow A$
$\alpha/K \rightsquigarrow B$	q	✕	r	q/r	No influence
$\alpha/K \rightsquigarrow A \wedge B$	s	✓	t	s/t	$\alpha/K \Diamond\!\!\rightsquigarrow A \wedge B$

Table 5.7a illustrates the noiseless data case modelling $\rightsquigarrow$. Since each line of the table is drawn from the same data set, the cases where α/K has been observed we may make some statements about the evidence tallies. If α/K has been observed n times and $\alpha/K \rightsquigarrow A$ then the positive evidence column for this line of the table will be n. Because B is constant every observation of A will also be an observation of B and since

the α/K data set has n elements the positive evidence count for $\alpha/K \leadsto B$ will also be n. By extension the positive evidence count for $\alpha/K \leadsto A \wedge B$ will also be n. Our only concern with negative evidence is that there be at least one occurrence and that this occurrence be associated with $\alpha/\neg K$. We see a tick in the negative evidence columns for $\alpha/K \leadsto A$ and $\alpha/K \leadsto A \wedge B$, this indicates an unspecified number of occurrences but there is at least one in each case. The negative evidence count for $\alpha/K \leadsto B$ is zero. With this evidence we see that $\alpha/K \leadsto A$ holds and that $\alpha/K \leadsto B$ is falsified. We may generate the negative evidence for $\alpha/K \leadsto A \wedge B$ from the negative evidence from the rows above. We have evidence of at least one instance of $\neg A$ and because B is constant this means that there is at least one instance of $\neg(A \wedge B)$ which provides negative evidence for $\alpha/K \leadsto A \wedge B$ indicating that the statement holds. The left hand side of equation 5.12 holds, the first expression on the right hand side also holds but the second is falsified providing a counterexample.

Moving to the evidence of table 5.7b we consider equation 5.12 in a noisy data environment. Despite the potential masking and uncertainty introduced by noise the fact that B is constant will not be obscured and it may be treated in the same manner as $\top$ in our consideration of $\overline{N}$. We see a 3 in the negative evidence column for $\alpha/K \leadsto A \wedge B$ indicating lack of influence and this falsifies the second expression on the right hand side of equation 5.12.

5.2.7 Syllogism, *S*

Chellas [30, page 271] lists this principle of syllogism as:

$$(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C)) \tag{5.13}$$

Casting this against our notion of influence gives:

$$(\alpha/K \leadsto A) \rightarrow ((A \rightarrow C) \rightarrow (\alpha/K \leadsto C)) \tag{5.14}$$

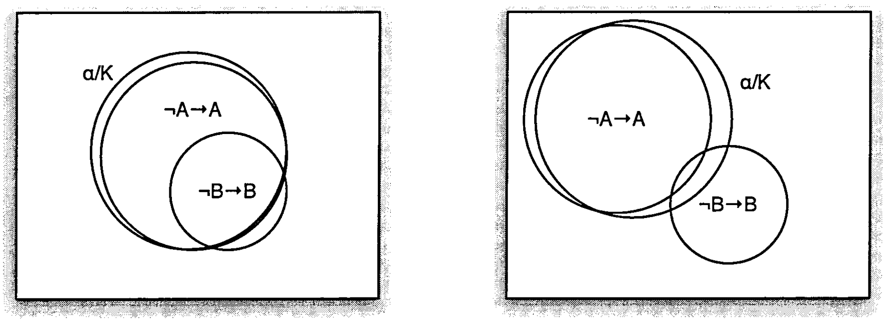
This rule requires careful consideration as it has bearings on coaching agents abilities to infer cause and effect and to divine necessary preconditions for actions. There are two ways that propositions may be linked; they may be inextricably linked so that the value of B mirrors that of A or they may be linked by transition so that a change from $\neg A$ to A causes B to hold.

Since A and B may have different truth values this is not a case of equivalent propositions, which are handled by *RE*, it is a case of sometimes subtle interaction between elements of the agent world. If an agent is carrying a bucket and we take A as meaning *bucket empty* and consider B as *no fire burning*. For this reading of B to be of note then it must have been the result of a change in the environment so there is an implicit transition of $\neg A$ to A involved in the proposition C. Both propositions are independent at the start of a cycle and the link between A and B is *strong*. We attempt to falsify this in table 5.8 by illustrating a scenario where an agent drives a transition from $\neg A$ to A and where this happens coincidentally with $\neg C$ For a counterexample to hold we would expect to see evidence values supporting figure 5.2(b), $q/r >> s/t$.

Table 5.8: Evidence and S

Hypothesis	Evidence			$P{:}C$ ratio	Conclusion
	Positive-	Negative-	Counter-		
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K \Diamond\!\!\rightsquigarrow A$
$\alpha/\neg K \rightsquigarrow B$	q	✓	r	q/r	$\alpha/\neg K \Diamond\!\!\rightsquigarrow B$
$\alpha/K \rightsquigarrow B$	s	✓	t	s/t	$\alpha/K \Diamond\!\!\rightsquigarrow B$

However, if A and B are relates as described above then coaching agents will observe evidence supporting figure 5.2(a) with s/t and n/p supporting their relevant hypotheses and, assuming that no other action in α's choice partition influences A in the same way as K, q/r will be small relative to the evidence supporting other hypotheses. The $P{:}C$ ratios for the first and third hypotheses in table 5.8 will be supportive because there are no other agents involved in this influence, α is self contained provided that $\neg A$ holds at the start of a cycle.



(a) $A \supset C$ (b) $A \not\supset C$

Figure 5.2: Evidence for influence (a) and lack of influence (b) in a noisy environment

This leads to a dilemma, a coach may have two hypotheses with the same set of precepts. Which does it seed? This is discussed in chapter 6 and S supports simply seeding the more influential hypothesis and this is intuitively correct as both hypotheses involve the same agent choice.

5.2.8 Tautology, T

Chellas [30, page 6] lists T as:

$$\Box A \rightarrow A \tag{5.15}$$

T states that if something is necessary then that something is the case. If α sees to it that A holds then A must hold.

$$T.\ [\alpha \leadsto : A] \supset A \tag{5.16}$$

If $\alpha/K \leadsto A$ holds then we will be presented with evidence like that of the top hypothesis line of table 5.9. This has positive evidence, negative evidence and no counter evidence. From our earlier discussion on evidence we know that if counter evidence is presented along with positive and negative evidence then it

Table 5.9: Example and counterexamples for T

Hypothesis	Evidence			$P{:}C$ ratio	Conclusion
	Positive-	Negative-	Counter-		
$\alpha/K \rightsquigarrow A$	n	✓	0	?	$\alpha/K \rightsquigarrow A$
$\alpha/K \rightsquigarrow A$	p	✓	q	?	Other influence
$\alpha/K \rightsquigarrow A$	n	×	0	?	No influence

is indicative of another influencing factor. The actor specification in equation 5.16 will then be incomplete falsifying the expression. This is illustrated in the second hypothesis of table 5.9. It may also be that A is constant and that α has no influence. Were this the case then there would be no counterexamples with $\neg A$ following α/K and there would be no negative evidence and equation 5.16 would be falsified. For 5.16 to hold it must be that case that A follows $[\alpha \rightsquigarrow: A]$.

5.2.9 What is necessarily so is necessarily necessarily so, 4

4 presents difficulties, Horty and Belnap [67] indicate that it makes a considerable claim about agency, an agent that sees to it that A holds also sees to it that it sees to it that A holds. Chellas lists this as:

$$4.\ \Box A \rightarrow \Box\Box A \tag{5.17}$$

Some modal accounts of action deny 4, for example those of Brown [26] and Chellas [29], but this is a necessary aspect of our approach as it places a condition of independence on agents, α*'s* choice may be influenced by β*'s* actions but β has no *control* over α*'s* choice. Recall that we consider agents as choice partitions overlaid on a discrete branching time background. 4 implies that the partitioning of histories at a moment is fully under the control of the agent choosing at that moment. In a branching time framework 4 is supported by the fundamental assumption that the choices available to an agent at a moment may be represented by the partitioning of histories through that moment.

$$4.\ [\alpha \rightsquigarrow: A] \rightarrow [\alpha \rightsquigarrow: [\alpha \rightsquigarrow: A]] \tag{5.18}$$

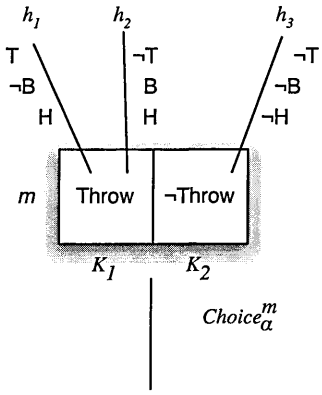


Figure 5.3: The choices available to a poor darts player

Recall Kenny's example of a poor darts player (see Kenny [72]) illustrated in figure 5.3 (also shown in section 2.7). This presupposes a noise free environment and it is difficult to deny that for H to hold the agent must select choice K_1 and throw a dart. In a noisy environment the dart may hit a dart already on the board or bounce off of a dividing wire leading to $\neg H$ following K_1. Viewing this in our hypothesis evidence context we may see evidence such as that of table 5.10a. This presupposes that the agent has only

Table 5.10: Evidence and 4

Hypothesis	Evidence			$P{:}C$ ratio	Conclusion
	Positive-	Negative-	Counter-		
(5.10a): agent has only two choices					
$\alpha/K_1 \rightsquigarrow H$	p	✓	q	p/q	$\alpha/K_1 \Diamond\!\!\rightsquigarrow H$
$\alpha/K_2 \rightsquigarrow H$	0	×	0	0	$\alpha/K_2 \not\rightsquigarrow H$
(5.10b): agent has multiple choices					
$\alpha/K \rightsquigarrow H$	p	✓	q	p/q	$\alpha/K \Diamond\!\!\rightsquigarrow H$
$\alpha/\neg K \rightsquigarrow H$	r	✓	s	r/s	$\neg(\alpha/K \Diamond\!\!\rightsquigarrow H)$
$\alpha/L \rightsquigarrow H$	t	✓	u	t/u	$\alpha/L \Diamond\!\!\rightsquigarrow H$

two actions available, one that has influence and another that does not. In this case we have no evidence that K_2 is influential and the agent must choose K_1 in order to bring about H. If the agent has more choices –

some of which may have influence over H – then we may see evidence such as that of table 5.10b which also supposes noise. Here we will see that $p/q > r/s$, the hypothesis $\alpha/K\Diamond\!\!\leadsto H$ has stronger supporting evidence. If the $\neg K$ binary choice partition contains another action, L that influences H then this will cause the evidence for $\alpha/\neg K \leadsto H$ to strengthen and will also cause the instantiation of another hypothesis, $\alpha/L \leadsto H$. Evidence for $\alpha/K \leadsto H$ and $\alpha/L \leadsto H$ will be greater than that for $\alpha/\neg K \leadsto H$ indicating that the agent has a choice in the matter and 4 holds.

5.2.10 4_{agent}

4 brought considerable claims about agency and the other agent extension makes further and stronger claims.

$$4_{agent}.\ [\alpha \leadsto: A] \rightarrow [\beta \leadsto: [\alpha \leadsto: A]] \tag{5.19}$$

Bringing the *other agent* into an action gives equation 5.19 which states that β has some influence over α seeing to it that A holds.

Table 5.11: Evidence and 4_{agent}

Hypothesis	Evidence			P:C ratio	Conclusion
	Positive-	Negative-	Counter-		
$\alpha/K \leadsto A$	p	✓	q	p/q	$\alpha/K\Diamond\!\!\leadsto A$
$\{\beta/L;\alpha/K\} \leadsto A$	r	✓	r	r/s	$\{\beta/L;\alpha/K\}\Diamond\!\!\leadsto A$
$\{\beta/L\Vert\alpha/K\} \leadsto A$	t	✓	u	t/u	$\{\beta/L\Vert\alpha/K\}\Diamond\!\!\leadsto A$

This is a rule that, obviously, may only be considered in cases where there is a joint influence requirement and will not be applied to cases where a coach is satisfied that a single agent has influence on A. 4_{agent} represents the nested STIT discussed at length by Belnap at al. [8, chapter 10] and the stit representation admits both serial and parallel action readings. If we combine aspects of Hansson's situationist logic, Hansson [57], with our notion of gateways introduced in section 3.2 and state that the partitioning is fully under an agent's control then we may view agent action in a deontic context. Agents have different *oughts* in different sets of circumstances. Hansson indicates that what an agent ought to do depends on its context and in a multi-agent behaviour α may open a gateway but it is β's choice whether or not to use that gateway. Returning to evidence, table 5.11 indicates what we may see. If α has single agent influence then p/q will indicate this and there will be no need for the coaching agent to consider two agent hypotheses, r/s and t/u will be lower indicating that the single agent hypothesis provides a better explanation for events.

4 on its own holds and this represents agency by indicating that deterministic aspects of the future may be mapped onto agent choices. 4_{agent} fails, if this were to hold then it would infer that some agents had control rather than influence over other agents.

5.2.11 Rule of equivalence *RE*

Chellas lists the rule of equivalence as:

$$RE.\ \frac{A \leftrightarrow B}{\Box A \leftrightarrow \Box B} \tag{5.20}$$

Casting this into influence based view gives:

$$RE.\ \frac{A \leftrightarrow B}{[\alpha \rightsquigarrow : A] \leftrightarrow [\alpha \rightsquigarrow : B]} \tag{5.21}$$

RE says that equivalent propositions are equally necessary. This relates to several aspects of agent behaviour in a coached environment. We must be careful with the notion of equivalence. If two propositions are absolutely equivalent, that is to say that they are the same but simply carry different labels or names, then one would expect to see evidence tallies matching exactly even in a noisy multi agent environment. If, however, the equivalence is both propositions are the result of the same choice then evidence tallies may not match exactly. In this case a coach would observe supporting two hypotheses, one that a given action leads to A and another that the same action leads to B.

Let us assume two hypotheses, one that $[\alpha \rightsquigarrow : A]$ and one that $\neg[\alpha \rightsquigarrow : B]$ for α/K. Let us also assume that α/K brings about A and brings about B

Table 5.12: Evidence supporting *RE*

Hypothesis	Evidence: Positive-	Evidence: Negative-	Evidence: Counter-	*P*:*C* ratio	Conclusion
(5.12a): $\rightsquigarrow$ example					
$\alpha/K \rightsquigarrow A$	n	✓	0	∞	$\alpha/K \rightsquigarrow A$
$\neg\alpha/K \rightsquigarrow B$	p	✓	q	p/n	$\neg(\neg\alpha/K \rightsquigarrow B)$
$\alpha/K \rightsquigarrow B$	n	✓	0	∞	$\alpha/K \rightsquigarrow B$
(5.12b): $\diamond\!\!\rightsquigarrow$ example					
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K \diamond\!\!\rightsquigarrow A$
$\neg\alpha/K \rightsquigarrow B$	q	✓	r	q/r	$\neg(\neg\alpha/K \diamond\!\!\rightsquigarrow B)$
$\alpha/K \rightsquigarrow B$	s	✓	t	s/t	$\alpha/K \diamond\!\!\rightsquigarrow B$

In the noiseless example of table 5.12a we see that the equivalence of A and B is reflected in the positive, negative and counter evidence tallies. Assuming that the $\alpha/K \rightsquigarrow A$ hypothesis holds, as in the statement above, and that A and B are equivalent propositions then there will be counter evidence for the $\neg\alpha/K \rightsquigarrow B$ hypothesis. The conclusion for the negative hypothesis, $\neg\alpha/K \rightsquigarrow B$, is that it does not hold.

The counter evidence which negates the $\neg\alpha/K \rightsquigarrow B$ hypothesis moves us to a noisy environment where we consider a may lead to result. Here the *P*:*C* ratio for $\neg\alpha/K \rightsquigarrow B$ will be significantly smaller than that for both $\alpha/K \rightsquigarrow A$ and $\alpha/K \rightsquigarrow B$. Note that n/p and s/t are not necessarily equal, other agents may play a part in observed evidence.

RE considers equivalent propositions from the point of view of a single actor as the agent of change. In a multi agent environment the ability to extend a hypothesis across groups of agents, agents which are members of some equivalence class, will allow a coach to develop behaviours applicable to a greater number of actor agents.

5.2.12 RE_{agent}

Before considering the agent extension to RE we address the question of *agent equivalence*, raised in section 5.2.11, which allows a coach to infer that one agent's ability to bring about A is transferable to other agents. Agent ability is driven by agent choice and we call similarly capable agents *choice class equivalent.*

Definition 40 *Given two agents, α and β, with choice sets $Choice_\alpha$ and $Choice_\beta$ respectively we say that α and β are choice class equivalent for a choice C iff $C \in \{Choice_\alpha \cap Choice_\beta\}$.*

Agents belonging to the same agent class will be choice class equivalent by default and the definition above may be extended across agent classes where these classes share common agent choices. This refinement of agent class equivalence to choice class equivalence removes a degree of coarseness from coach reasoning.

RE_{agent} maps *RE's* claim of the equivalence of propositions onto the domain of agents. This is a necessary mapping because coaching agents deal with the world in an abstract manner which treats agent / action pairs as propositions.

If two agents, α and β are members of the same agent class then it is trivially true that given the same local environment and reading *seeing to it* as an indicator of ability then α seeing to it that A holds is equivalent to β seeing to it that A holds. Our hypotheses are built on agent / action pairs and take the form $\alpha/K \rightsquigarrow A$. If agents are choice class equivalent for K then this represents an intersection of their agency and since agents may execute only one choice at any instant they are equivalent when constrained to choice equivalence.

$$RE_{agent}. \quad \frac{\alpha/K \leftrightarrow \beta/L}{[\alpha \rightsquigarrow: A] \leftrightarrow [\beta \rightsquigarrow: A]} \tag{5.22}$$

This is an important rule for our system, if RE_{agent} fails then a coaching agent can not assume that a good behaviour exhibited by one agent of a particular choice class may be transferred successfully to another agent of that class.

Table 5.13: Evidence supporting RE_{agent}

Hypothesis	Evidence			P:C ratio	Conclusion
	Positive-	Negative-	Counter-		
$\alpha/K \rightsquigarrow A$	n	✓	p	n/p	$\alpha/K\Diamond\!\!\rightsquigarrow A$
$\beta/L \rightsquigarrow A$	q	✓	r	q/r	$\beta/L\Diamond\!\!\rightsquigarrow A$

Given two agents, α and β, belonging to separate agent classes and two different actions, K and L for α and β respectively, a coach may see evidence such as that of table 5.13. Both α and β present evidence of being able to bring about A and in this case the coach will simply consider this an equivalent ability for each agent class and seed a behaviour for each class. Functionally both behaviours are equivalent, the coach is unable to see a difference and simply treats them as equivalent.

5.3 From characterisation to implementation

The partial characterisation presented in this chapter indicates that an evidence based approach to identifying agent influence will allow coaching agents to reason about chains of actions, equivalence of actions and equivalence of propositions in a modal context. It also indicates that the system will not support claims, such as those of C_{agent}, which may lead the system to attempt to reason through cases of mutual exclusion. Similarly 4_{agent} does not hold because this would carry implications of agents being able to see to it that other agents bring propositions about and this is much stronger than our claims for influence. We carry these properties over to the following chapter where we consider how a coaching agent may use them in managing evidence to divine agent ability.

Chapter 6

Coaching agents

We have developed a theory of influence, discussed two influence operators and have built a partial logical characterisation of these operators. This characterisation has two purposes, these are to explore the behaviour of influence and to provide guidance in implementing coaching agents within a system. Horty, Belnap, Xu and others concentrate on an abstract theory of STIT and ability, one that makes heavy demands on an agent's reasoning resources. There is literature on reasoning by resource bounded agents, Rubinstein [99] and Bratman [19] for example, but these carry a reasoning component which makes them suitable for agents with cognitive abilities. We intend to build coaching agents in the same manner as actor agents, as close to fully reactive as possible and with minimal cognitive ability. Here we outline coaching agent operation and the synthesis of behaviours for reactive agents.

6.1 The relationship between a coach and an agent

The roots of this research lie in emergent behaviour experiments with StarLogo (see Resnick [96]) which moved to C and further experiments in emergent behaviour [82]. Communications, in these settings, used the environment as an intermediary. In exploratory work by Logie et al. [82] the environment held pheromone like data which agents could read and write. Communications in this instance are more complex requiring that agents communicate more data than a simple pheromone level. We have mentioned behaviour patches and history patches. Briefly and informally, a history patch is a collection of agent data gathered over one operating cycle. It holds the agent's percepts before any action, the agent's action and its percepts following that action. A behaviour patch contains a set of weightings which bias an agent towards a single action and a set of precepts where coaching agents have observed this action as being influential.

The task of driving system evolution and learning falls on coaching agents. Note that we use the terms evolution and learning rather loosely. By evolution we mean that an actor agent's behaviour patterns may change over time but it's underlying stochastic ability remains fixed. By learning we mean that the system's performance, in terms of it maximising agent, influence improves over time. A small residual stochastic element in agent behaviour means that although the agents in the system may not reach optimal behaviour they will be able to adapt to changes in its environment. Coaching agents operate with a general hypothesis that agents may influence their environment and, by extension, each other. Coaching agents will analyse historical data with a view to locating this agent influence.

This chapter outlines how coaching agents will approach the analysis of agent behaviour history data and illustrates the link between hypotheses of influence, as outlined in section 4.4, and their appearance when represented in a branching time framework. This branching time representation will allow us to consider the hypotheses in terms of stit semantics leading to a notion of influence frames which acts as an intermediate step to our reading of nested other agent stit expressions.

Coaching agent operation is described in four steps before moving to a more general overview. The first step is to describe some aspects of the environment and indicate why coaching agents are necessary. The second step informally outlines how the coach operates and leads to the third step which explores some simple agent interactions so as to build the notion of an influence frame. The fourth step sees us taking this influence frame and characterises it more generally.

6.2 Preparing for the coaching operation

We have already outlined coaching agent operation as gathering agent history data from an environment, applying this data to an abstract agent representation, synthesising new agent behaviour based on these observed data and seeding the environment with these behaviours. We begin by stating that agent percepts are a true representation of their environment. Agents do not suffer from faulty or noisy data and if they operate in a noisy environment then their percepts will accurately represent this noise. We also state that the events or world states that an external observer deems "good", our system norms, are possible, this is a requirement of the *ought implies can* deontic identity. We state as suppositions:

- Suppose that there is a class of agents that has the ability to see to it that A. Note that this is an abstract agent class as a result of our treating agents as abstract choice mechanisms.
- Suppose that an agent α in that class is located so as to be able to exercise that influence.

It may be that an agent has several choices which enable it to bring about A, this is implicit in statements such as $[\alpha\ istit\colon A]$ which make no reference beyond an agent identity. Leaving such things implicit is potentially dangerous. Different choices may bring about A but they may also have different preconditions so they can not be considered as equivalent choices when synthesising behaviours. Coaching agents will consider actor agents as abstract choice classes (definition 36) and this allows them to analyse influence at a choice level rather than at a coarser agent level. As an intermediate step we extend standard STIT notation to accommodate this finer granularity by indexing STIT operations against a particular agent choice.

Definition 41 *A statement of the form* $[\alpha\ istit_K\colon A]$ *says that by choosing K from its choice set agent* α *is able to instantaneously bring about A. Note that K may be a choice to refrain from acting as well as a choice to act.*

We suppose that a coaching agent hypothesises that one agent, β, has an influence of α*'s* abilities and characterise the coach agent's operation by the following four steps with ISTIT representing a choice by which the agent immediately brings about A.

1. The coaching agent sees partial histories generated by α which contain transitions from $\neg A$ to A.
2. The coaching agent will have a collection of partial histories where α chooses a particular action, K and $[\alpha\ istit_K\colon A]$ holds. The coach may also see histories where that action is selected and $[\alpha\ istit_K\colon A]$ does not hold.

3. The coaching agent then examines histories for evidence of actions by other agents on histories where $[\alpha\ istit_K\colon A]$ holds which are absent on histories where $[\alpha\ istit\colon A]$ does not hold.

4. The coaching agent aggregates these data and attempts to build an influence frame that models $[\beta\ istit\colon [\alpha\ istit_K\colon A]]$

Before proceeding we expand some of the notions in the list above. Item 1 above specifies *partial histories*, because agents in this system are resource bounded and are only able to record brief histories of percepts and actions. Coaching agents deal with partial histories, a record of an agent's most recent precepts, actions and postcepts.

Definition 42 *A partial history is an arbitrarily bounded history from an agent's situated perspective and contains an ordered list of an agent's most recent precepts, postcepts and actions.*

Item 1 states that coaches are looking for histories where transitions from $\neg A$ to A occur. This does not necessarily imply that agent influence is present. A may spontaneously occur and just as spontaneously clear, an agent may record the transition causing it to appear in its history but it may not have had any influence on observed transition of S. Certain elements of the environment may be cyclic and while they may respond to agent intervention the response may be variable and depend on the timing of the agent's action. Where an agent has direct and unambiguous influence over something then changes may appear to be consistent with an action or sequence of actions.

Definition 43 *A state change in the environment, a transition from $\neg A$ to A, indicates that an agent has perceived the change but does not imply that an agent has had any influence on the change.*

We informally introduce the notion of an influence frame as a two bounding instants. A choice at the earlier instant makes A possible by some choice at a later instant. This witnessing instant at the start of an influence frame is illustrated in figure 6.1 where α/K at m makes A possible following some future agent choice m'. When a nested influence is *framed*, as indicated in item 3, we refer to the first agent, β in this case, as the *outer* agent and α in the example as the *inner* agent. On entry to an influence frame an agent may not have exercised any influence proper but it has exercised *potential influence* over A.

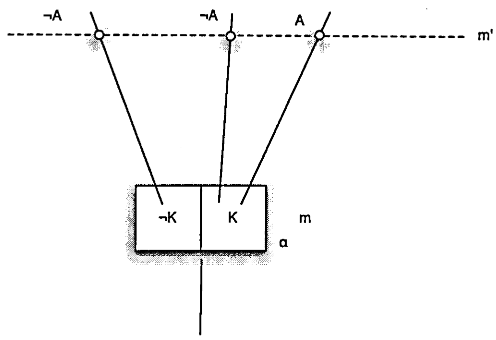


Figure 6.1: STIT frame with A-potential

Definition 44 *Let $H_A(m)$ be those histories for which $h \in H_A(m) \rightarrow m/h \in v(A)$. A history has* A-potential influence *at moment m when $H_A(m) \neq \emptyset \neq H_{\neg A}(m)$*

Intuitively, at this witnessing moment, w an agent acts in some way that makes A possible within what we term an *influence frame*. The witness moment also has a negative condition, since we are dealing with the possibility here the positive condition is that an action leads to A being necessarily possible and the negative condition leads to A being necessarily not possible and this is illustrated in figure 6.2. This simple notion

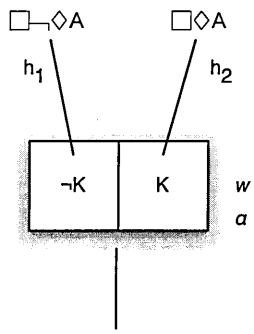


Figure 6.2: A moment where α is witness to $\Diamond A$

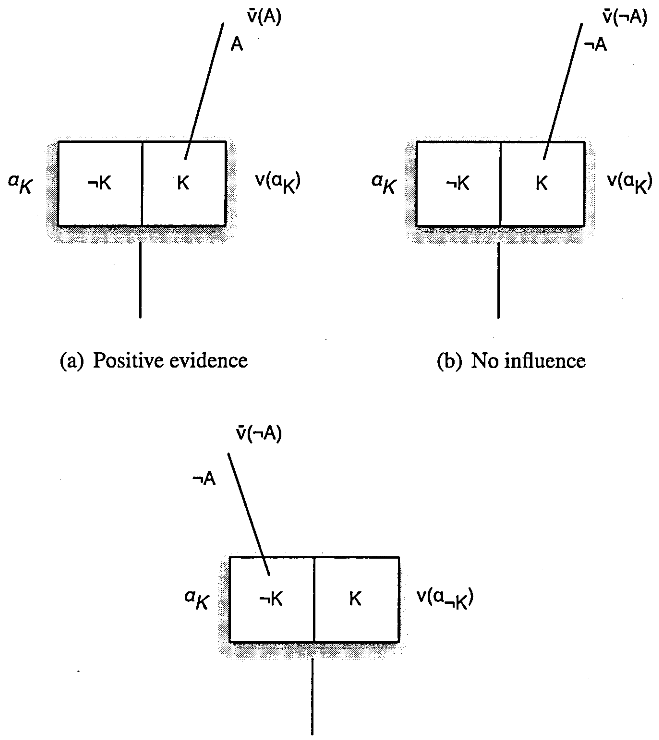
has several interesting properties which we shall outline before a formal definition. The single witnessing moment does not mean that we are constraining the frame to a representation of the influence of a single entity. The choice at the witnessing moment may be that of a single agent or it may be a group of agents working together serially or in parallel. The choice which brings A about at the later evaluation instant may be made by another agent and we must be clear about which agents which we are admitting to the influence frame. If, in figure 6.2, the future follows h_1 then there may be some instant before the evaluation instant where α selects choice K meaning that the negative condition potentially does not hold. Consider the presence of another agent, if that agent were to intervene and select action K then it may independently bring about a state – from the negative branch – where A may be brought about. To allow for this we state that the negative condition is only required to hold at $w + 1$, instantaneously after the choice at the witness moment.

6.3 Characterising evidence of immediate influence

Influence manifests itself in two ways. Assume an action, K, which an agent may or may not execute. Assume, also, a proposition, A, relating to the agent's world. We hypothesise that an agent, α, has influence on A and that by choosing K, α may bring it about that A holds.

The partial histories in figure 6.3 illustrate three possible outcomes of α's choice $\{\neg K, K\}$ at a moment. We have two valuation functions, v for α's choice and $\overline{v}$ for a postcept of proposition, A, which we assume that α may influence. Figure 6.3(a) indicates positive evidence that a choice of K by α influences A being brought about. On its own this is insufficient to support any hypothesis that α's choice of K has influence because it only partially satisfies the conditions discussed in section 4.4. Seeing instances of 6.3(a) in addition to instances of 6.3(c), indicates that if K is not selected then A is not brought about, we may infer that α does have some influence over A. A counterexample to influence is illustrated in figure 6.3(b) where action K does not bring about A. Coaching agents may encounter some or all of these scenarios and it is the combination of these evidence features that will guide the coach to search for other agent influence and nested behaviours.

If we characterise influence in terms of an istit then we may say that $[\alpha\ istit_k\colon A]$ (so $\mathcal{M}, m/h \models [\alpha\ istit_k\colon A]\ \forall\ m/h$). We see, also, that the valuations v and $\overline{v}$ are related:



(a) Positive evidence

(b) No influence

(c) Influence

Figure 6.3: The components of influence (a,c) and lack of influence (b) of α/K over A

a) As a postcept we see $\overline{v}(A) = v(\alpha_K)$

b) As a postcept we see $\overline{v}(\neg A) \cap v(\alpha_{\neg K}) \neq \varnothing$

If $\neg[\alpha\ istit_k\colon A]$ holds then we will see one of the following possible sets of circumstances.

1) i) $\overline{v}(\neg A) \cap v(\alpha_K) \neq \varnothing$

or

ii) $\overline{v}(\neg A) \cap v(\alpha_K) = \varnothing$

or both.

2) i) $(\overline{v}(A) \cap v(\alpha_K) \neq \emptyset) \;\cap\; (\overline{v}(\neg A) \cap v(\alpha_{\neg K}) \neq \emptyset)$

ii) $\overline{v}(A) \cap v(\alpha_K) = \emptyset$

6.3.1 Parallel influence

Suppose, also, that there is a β such that $\neg[\beta \ istit_{K'} : A]$ and that 2-i, from the list above, holds for β. Suppose, in addition, that $\overline{v}(A) \cap v(\beta_{K'}) \neq \emptyset$ so β is not prevented from having influence.

What if $\overline{v}(A) = v(\alpha_K) \cap v(\beta_{K'})$

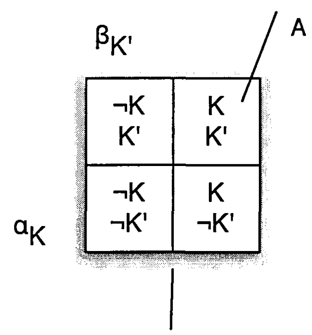


Figure 6.4: α and β's joint ability at a single moment

Then we say $[\alpha \| \beta \ istit_{K,K'} : A]$ but $\neg[\alpha \ istit_K : A]$ and $\neg[\beta \ istit_{K'} : A]$

6.3.2 Serial influence

Suppose that we observe consistent evidence of α's having influence on A, as in figure 6.3(c) but we also observe both positive evidence, figure 6.3(a), and evidence of no influence, 6.3(b). We may hypothesise that although α is the final actor that brings about A there is also some other influence playing a part, an influence outside of α's control. We assume that this other influence is some other agent, β, and that β's earlier choices have some influence on the outcome of α's choices. Intuitively such serial influence is not simply a chaining of two instances of single agent influence. Our analysis of agent influence is based on choice and, consequently, agency. If β does indeed have influence over the results of α's choices then that will be evident only when α exercises its choices appropriately.

We assume that β, which acts before α has two choices, K' and $\neg K'$, α has two choices, K and $\neg K$. We wish to show $[\beta ; \alpha\ istit_{K',K} : A]$ which says that β / K' followed by α / K can jointly and instantaneously see to it that A holds. Before proceeding let us consider the meaning of immediately in this context. Immediately is intuitively obvious in the single agent and joint parallel cases, actions are taken so as to bring about a result at the next instant. This is clearly not possible when one agent's action follows another and actions occur on integral time ticks and we consider this in the following sections.

6.4 Nested stit expressions

Horty and Belnap [67] note that STIT operations encourage nesting and in a single agent setting nesting may be used to analyse concepts such as refraining from acting. For example an agent refraining from is seeing to it that A may be said to be seeing to it that it does not see to it that A, $[\alpha\ stit\colon \neg[\alpha\ stit\colon A]]$. A nested STIT expression describes sequences where the sentence that is the final object of the statement may depend on a number of actions. For example, if someone wishes to check an email account then they must make sure that their computer is turned on. Rephrasing this with a STIT element, in order to see to it that they check email a computer user must see to it that the machine is turned on. Lorini et al. [84] indicate that stit is an S5 modal operator and Broersen [21] notes that nested S5 operators may be replaced by logically equivalent non-nested formulas. Our interest in nesting lies in so-called *other agent* constructions and the difficulties that they bring as discussed in section 3.3.

We assume than an action K by α in some circumstances brings about A and in some circumstances brings about $\neg A$ and that this is dependent of β's action. When β acts so as to bring the system into the inner section of the nested statement then whenever α executes K then A will be the result.

When the expression enters the nested braces the fact that β has exerted its influence means that A will be brought about when α chooses K. α retains its agency and has its influence extended by gaining the ability to influence A as a result of β's action. We refer to the agents as outer and inner agents depending on their role in the nested statement.

In simple cases with no intermediate moments between the witnessing and final action moments the role of the inner agent is clear. Maintaining agency means that the inner agent must have a choice. Figure 6.5 shows a simple situation where the outer agent's action at m_w satisfies a negative condition, $\neg A$ at h_1 and gives the inner agent a choice at m_1 where it is able to select a history leading to either A or $\neg A$. A negative condition is necessary and a positive condition must be possible on another branch of the history tree. Figure

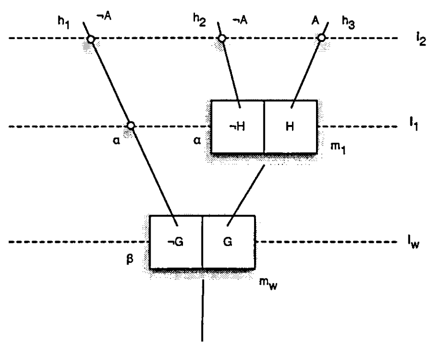


Figure 6.5: Branching time frame where an other-agent nested STIT holds

6.6 is similar in that its witness moment satisfies the negative condition requirement but the other branch is always positive. The inner agent, in this case, has no choice and is unable to exercise its agency so this does not constitute a nested STIT as agency does not hold throughout. The notion of settledness cannot extend from the outer agent to the object of the nested statement as this would remove agency from the inner agent.

If there are intermediate moments between the witnessing and final actions then things become a little more complex. It may be possible for an inner agent to have no choice as a result of its own actions.

Figure 6.7 illustrates a case where at m_1 α has a choice, D that leads to $\neg A$ but at the subsequent m_2 α has no choice, both $\neg B$ and B lead to histories where A holds. This satisfies the inner agency requirement because the lack of choice at m_2 is a result of a choice by α and is under the domain of α's agency.

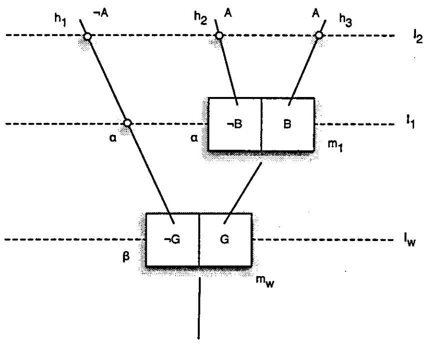


Figure 6.6: Branching time frame where an other-agent nested STIT does not hold

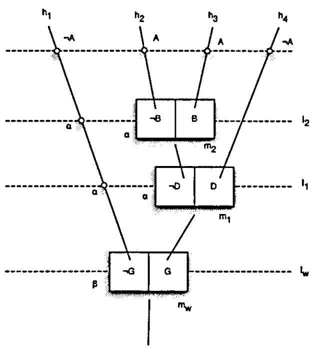


Figure 6.7: Branching time frame where nested STIT holds despite vacuous choice at I_2

Our early intuition was that the evaluation of an other agent nested influence statement has two frames, an outer frame defining the witnessing action and an inner frame defining the completion action. Viewing the agents together as a system then the system enters the inner frame when the outer agent acts in such a way as to instantaneously guarantee that the other agent may bring about A. There is a negative condition on this that if the outer agent may also act in such a way that makes it not possible for the inner agent to bring about A. This, however, implies that the outer agent's influence extends along the negative branch at least as far as the evaluation instant and this is not correct.

The inner frame, which is entered as a result of a choice by the outer agent at an earlier moment, has at least one moment where a choice can guarantee A and at least one moment where it can guarantee $\neg A$ at the same evaluation instant. Figure 6.8 indicates where α has no agency.

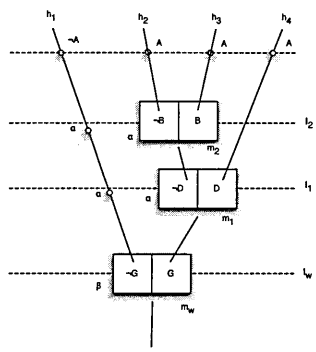


Figure 6.8: Branching time frame where nested STIT does not holds because all contained choices are vacuous

6.5 An exploration of two agent interaction

The examples above outline our notion of how nested statements should appear when framed in a single, multi agent STIT frame. A small program was used to generate random traces allowing us to explore two agent behaviour. It was assumed that the two agents were bound together but that they chose their actions independently. The agents were characterised as a carpenter, β and an apprentice, α. Each agent had five actions with each action having equal probability. Carpenter agents had *hit*, $\neg$*hit* and three other actions. Apprentice agents had one *give* action with four other actions. It was assumed that the carpenter agent started without a hammer and when α *gives* a hammer (which, as before, we refer to as a token) then it has persistence and β may use it for every following *hit*. When α hands a token to β the token becomes available for use in the next cycle. If the token is used to hit a nail, N, then the nail appears driven at the end of that agent cycle.

A sample of data generated is shown in table 6.1 where X for both the α and β agents means *other* action, H is hit, G is give and A is the driven state of the nail. The top row lists the most recent events with earlier events are in sequence below.

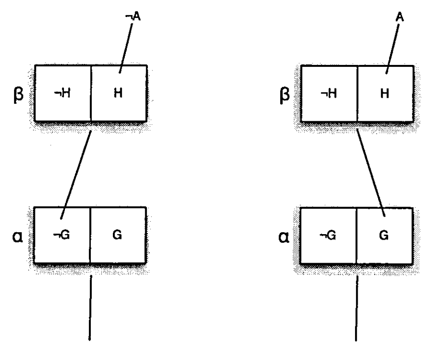
Table 6.1: Behaviour histories for carpenter, β, and apprentice, α, agents.

β	α	N	β	α	N	β	α	N	β	α	N	β	α	N
$\neg H$	X	$\neg A$	H	G	A	$\neg H$	X	$\neg A$	X	X	$\neg A$	X	G	$\neg A$
X	G	$\neg A$	H	X	A	X	X	$\neg A$	$\neg H$	X	$\neg A$	$\neg H$	X	$\neg A$
X	X	$\neg A$	$\neg H$	X	$\neg A$	X	X	$\neg A$	$\neg H$	G	$\neg A$	X	X	$\neg A$
H	X	$\neg A$	X	X	$\neg A$	X	X	$\neg A$	X	X	$\neg A$	X	X	$\neg A$
X	X	$\neg A$	$\neg H$	G	$\neg A$	X	X	$\neg A$	X	G	$\neg A$	X	X	$\neg A$
X	X	$\neg A$	$\neg H$	X	$\neg A$	X	X	$\neg A$	X	G	$\neg A$	X	X	$\neg A$
X	X	$\neg A$	H	X	A	$\neg H$	X	$\neg A$	X	X	$\neg A$	$\neg H$	X	$\neg A$
X	X	$\neg A$	X	X	$\neg A$	$\neg H$	X	$\neg A$	X	G	$\neg A$	$\neg H$	X	$\neg A$
X	X	$\neg A$	X	G	$\neg A$	X	X	$\neg A$	H	X	A	X	X	$\neg A$
X	X	$\neg A$	$\neg H$	G	$\neg A$	H	G	$\neg A$	X	G	$\neg A$	X	X	$\neg A$

The coaching agent procedure begins by assembling agent classes, this is simply a case of the coach examining an agents histories and building a set of reported actions. This is done on the assumption that agents are autonomous and choose their actions. The coach represents these agent classes as choice partitions which it may superimpose on a STIT frame. The carpenter agent, for example, has a number of actions, $\{H, \neg H, \ldots\}$, the history data show that the H action occasionally produces A so we naively partition the carpenter's choice set into H and $\neg H$. We adopt a similar aggregation approach with the apprentice agent (this is a simplifying approach for the early stages).

From table 6.1 a coaching agent may see the carpenter's H action on single history trails as 6.9(a) and 6.9(b).

The coach may aggregate these to give figure 6.10 where we have labelled instants, for our convenience as I_w the witnessing instant, I_e the evaluation instant with two intermediate instants I_{i1} and I_{i2}.



(a) *hit* and $\neg G$ (b) *hit* and G

Figure 6.9: Possible histories for a joint sequential action

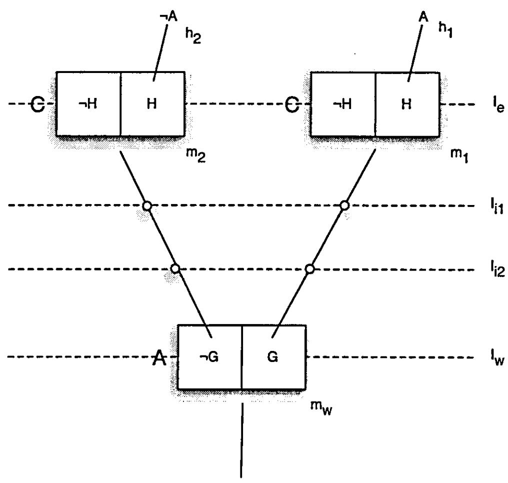


Figure 6.10: Branching time frame showing potential ISTIT action

6.6 An operational overview of a coaching agent

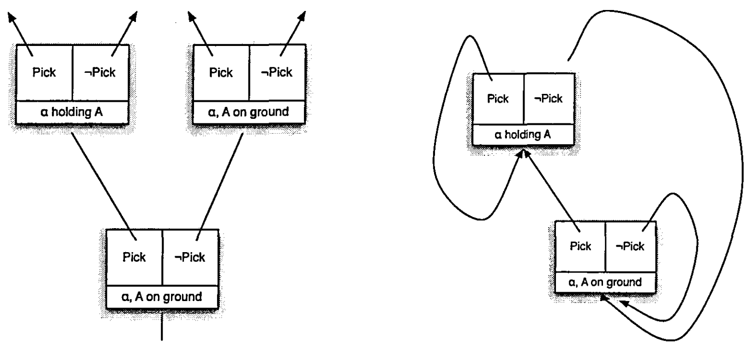
The coaching agent's main task is to randomly sample history data in the environment, use that data to generate or update hypotheses which are used to generate and seed new behaviours in the environment. We have hinted, in our intuitive outline of influence in section 3.2, that the coaching agent has a view of the environment that lies between the bounded actor agent view and a global observer's view. We now explore how coaching agents operate and in doing so we add detail to the notion of the coaching agent's intermediate view. The coaching agent does not have abilities to manipulate objects in the environment, its only way to drive change is to manage the behaviours of actor agents.

When a coach picks up a patch it examines it for evidence of change, our theory is based on influence and we consider change as prima facie evidence. If the history patch indicates change then the coach will check it against a hypothesis list which forms the core of its internal database. History patches provide two types of evidence, that directly related to a hypothesis and what we term *collateral evidence* where data may be used to toggle the negative condition of non related hypotheses. Direct hypotheses are indexed via agent class and agent action. For example, if a patch indicates that $\alpha/K \leadsto A$ then related hypotheses will be those where α/K holds so that the patch will provide positive evidence for a hypothesis that $\alpha/K \leadsto A$ and will provide counter evidence for other α/K hypotheses, for example $\alpha/K \leadsto B$. If a matching hypothesis is not present then the coach generates and stores a new hypothesis. If there is an extant matching hypothesis then the coach will update evidence for this hypothesis. Note that a single history patch may generate multiple hypotheses if there is evidence of multiple changes in the precept and postcept data.

After dealing with related hypotheses the coach will then use the data as collateral evidence, hypotheses are based on atomic propositions

6.6.1 Computational tractability, temporal evolution and coaching agents

We have described a tree like partial order, section 2.6.3, governing temporal aspects of a standard branching time structure. We have also described a time framework, section 4.1, which removes the density problem by discretising the intervals between instants. Such an ordering is useful for representing sections of a tree but its forward branching remains unbounded and this embodies some of the difficulties outlined in section 2.11. A standard branching time diagram may be thought of as a directed acyclic graph, potential cycles are eliminated by the relentless forward progression of time causes the repetition of world states at different points on the tree. Clearly this presents difficulties to coaching agents (and, indeed, to any bounded agent), trying to represent a complete temporal evolution of the world is going to generate large trees which will continue to grow as more data are acquired. Coaching agents certainly have an interest in temporal aspects of the world but they are more concerned with the transitions from state to state. By removing the strict temporal requirement from the ordering and reading transitions, informally, as *from starting state X action A brings about state Y* we may introduce cycles into the representation of the world.



(a) Standard branching time.

(b) Cyclic state graph

Figure 6.11: Repeated states in standard branching time and in a cyclic state graph

Consider a very simple world, one location, one object and one agent. The agent has two possible actions, *pick* and *drop* (we will view this as pick and $\neg$*pick* so as to remain consistent to our earlier representation of agent choice). In standard branching time a representation of the world continues to branch in an unbounded manner as in figure 6.11(a). By admitting cycles we condense the tree closing it to give the directed graph of figure 6.11(b). This implicitly represents the discrete branching time representation, the arcs in the directed graph represent an instantaneous STIT and have an implicit time interval. This extremely simple case presents an extreme where admitting cycles reduces a potentially unbounded tree to a graph with two nodes and four arcs. A world of sufficient complexity to offer any interest is going to present a much more complex graph but each loop in that graph represents a potentially large reduction in resource usage over a long sequence of operating cycles.

6.6.2 When a coach picks up a history patch...

Coaching agents are concerned with finding and maximising influence and our earlier discussion of hypotheses in section 4.4 was founded on the notion of agent / action combinations. We carry this into the coaching agent's internal data structures and adopt an action / negation structure for mapping the results of agent actions.

When a coach begins it has an empty database and it moves through its environment picking up history patches. At this stage if these history patches do not exhibit any evidence of influence then they are of no interest. When a coach encounters a history patch that indicates influence it will construct a hypothesis of the form $\alpha/K \rightsquigarrow S_1$ (where S_1 is a new world state) to describe its observation. The hypothesis specifies only the end state but the coaching agent's database will, as in the informal transition reading above, hold the start state for each transition and this start state will be implicitly held in any behaviour specification as a trigger state.

We shall be considering differences between sets of percepts so we define how these are identified.

Definition 45 *Given two sets of agent percepts S_0 before an action and S_1 after an action we define the change between these states $Changes_{S_0 \to S_1}$ as the complement of $S_0 \cap S_1$.*

$Changes_{S_0 \to S_1}$ is a set of atomic propositions and does not provide a firm foundation for building hypotheses, it may contain elements which are changed by influences other than the one under consideration. We constrain our hypotheses definitions by dividing the change set so that the hypothesis target is an atomic proposition.

Definition 46 *Given an agent, α, an action, K, a set of precepts S_0 and a set of postcepts S_1, if $Changes_{S_0 \to S_1} \neq \{\}$ then we may generate a set of hypotheses, $\{A \in Changes_{S_0 \to S_1} : \alpha/K \rightsquigarrow A\}$.*

If there are multiple changes then this may lead to a multiple hypotheses being generated.

Definition 47 *Given an atomic proposition A, an agent α and an action K we say that all cases where α/K brings about A are hypothesis equivalent.*

We consider a coaching agent which gathers patches indicating that an agent, we shall use a single agent class for the time being, has had influence with action K, the coaching agent may build a database

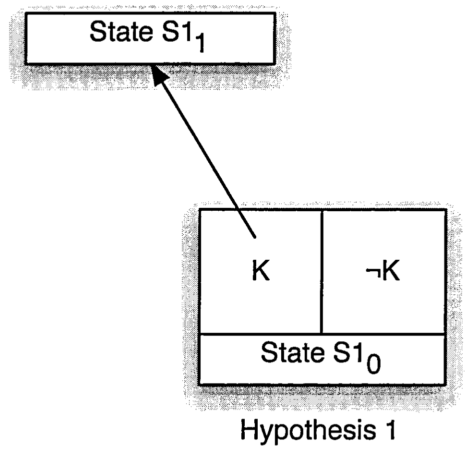
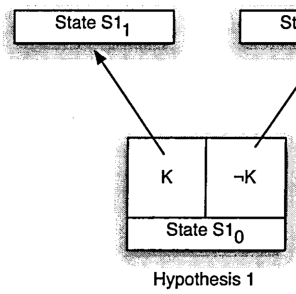


Figure 6.12: Coach database, hypothesis instantiation

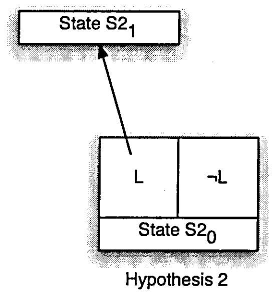
such as that of figure 6.12. The new hypothesis has a starting state, Sn_0, and a single next state, Sn_1. The time based partial order of the standard branching time structure is now represented as a following action step.

Although the coach has gathered data indicating that a change followed an agent action and has begun building a world database it does not have sufficient information to consider the hypothesis as fully formed. The transition data gathered so far represent only positive evidence. As more history patches are gathered the coach examines them against extant hypothesis structures. It may be that another history patch has the same (or similar) starting state from where the agent chooses another action and this results in another state. If, following the other action, the new state differs from $S1_1$ (in figure 6.12) in respect of the influence expressed in the transition from $S1_0$ to $S1_1$ then this satisfies the negative evidence requirement and allows the completion of the hypothesis as in figure 6.13(a). Note that the $\neg K$ resultant states are not recorded as part of the α/K hypothesis. Our hypothesis theory requires only one instance of negative evidence and that has been presented here so we have no further interest in the $\neg K$ branch from state $S1_0$.

If the new history patch has the same precepts and the agent executes a choice that is not represented elsewhere in the coach database and this patch shows some form of influence then it will also be used to



(a) Hypothesis 1 negative evidence



(b) Hypothesis 2

Figure 6.13: Coach database, collateral negative evidence

instantiate a new hypothesis, figure 6.13(b). As far as the α/K hypothesis family is concerned the right hand branch in figure 6.13(a), $\neg K$ choice at state $S1_0$, is complete. The $S1_0$ node has two forwards branches, the left branch represents α/K and the right branch represents $\alpha/\neg K$. The next states reachable along the right hand branch are explored more accurately by other hypotheses so our requirement of only a single instance of negative evidence does not reduce the potential state space coverage of the coaching agent's database.

So far our representation admits positive evidence and negative evidence but we have still to address counter evidence and evidence tallies. A hypothesis is built on the observed results of an agent's action, we need only see a single branch forwards from the negation of that action but we must follow each branch forwards from the positive action side.

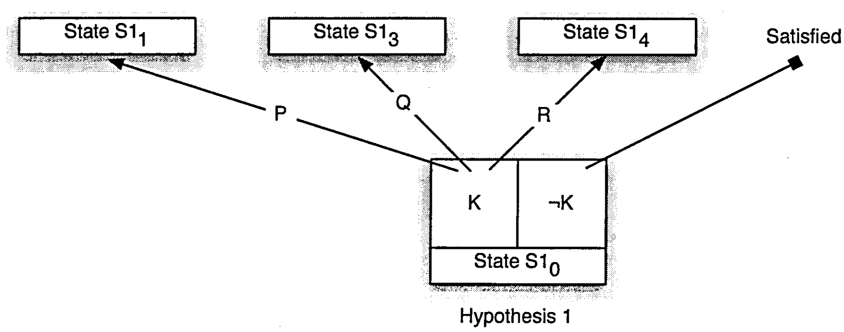


Figure 6.14: Coach database, multiple possible states following agent choosing K

Figure 6.14 illustrates this, the coach has collected other history patches where the precepts match $S1_0$ and the agent has chosen action K but the resulting state has not matched $S1_1$ (note that this is in terms of the atomic proposition of definition 46). We could explicitly represent this, as in figure 6.14, with the tallies for each state represented on the arcs from α/K at $S1_0$. This is an unnecessary complication and dilutes the focus of our foundation hypothesis. Instead we aggregate the $\neg S1_1$ states and we maintain a tally of these $\neg S1_1$ instances to allow its use in heuristics, such as those discussed in section 5.1, for hypothesis strength. This gives a database entry, for this single hypothesis, which resembles figure 6.15.

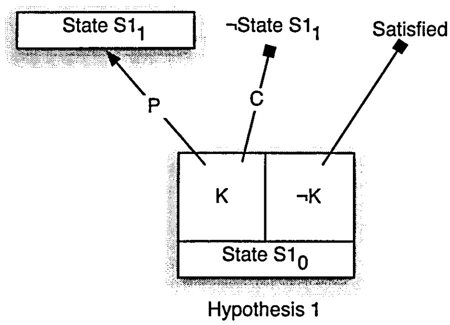


Figure 6.15: Coach database, representation of a binary hypothesis

6.6.3 Managing multiple hypotheses

On the route to a single hypothesis structure we have discarded certain transitions from our growing hypothesis state diagram. Given a history patch where an agent with precepts equivalent to state $S1_0$ (the states are numbered by their associated hypothesis) executes an L action leading to a next state $S2_1$ which is not hypothesis equivalent (definition 47) to $S1_1$ then this not only provides negative evidence for hypothesis 1 but it also allows the instantiation of a new hypothesis. The coach now has two hypotheses based on equivalent precepts, hypothesis 1 where $\alpha/K \rightsquigarrow S1_1$ and hypothesis 2 where $\alpha/L \rightsquigarrow S2_1$. Because these hypotheses have equivalent precepts they are joined by state equivalence links. This is illustrated in figure 6.16, note that the states $S1_0$ and $S2_0$ are *database states* and although they are equivalent as far as their precept content is

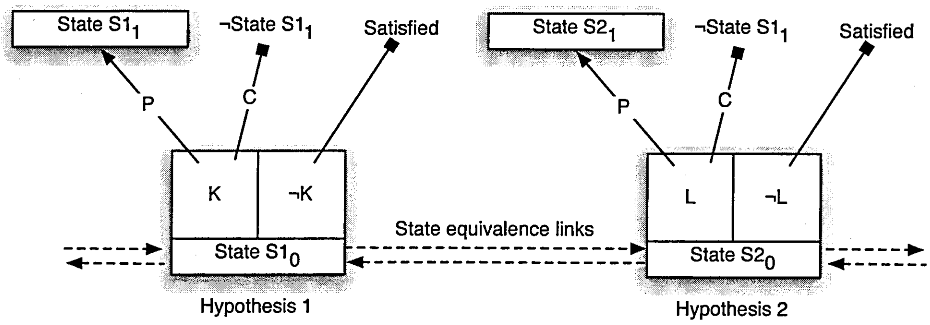


Figure 6.16: Coach database, multiple hypotheses on a state equivalence chain

concerned they differ in the action attached to the precepts. The database represents hypotheses rather than raw world states. This may build a chain of possible routes forward from a given state and each hypothesis along that chain will have positive and counter evidence tallies which may be used by heuristics to gauge its strength. The postcept states may be linked into other state equivalence chains so as to link to reachable hypotheses based on the following states.

This will give a structure similar to that of figure 6.17. The forwards branching of figure 6.17 is still unbounded. Since the partial order on states is based on agent actions rather than on time (although time is implicit) we may admit loops, as discussed above, allowing us to close off certain forward branches. If, for example, $S3_1$ and $S1_0$ are equivalent then α/M from $S3_0$ leads back to a preceding state closing a loop. This is illustrated in figure 6.18. Since agents have a limited repertoire of actions adopting a partial order based on actions rather than raw time will, intuitively, bound the graph of reachable world states.

6.6.4 Generating behaviour patches, maximising agent / action influence

As a coaching agent's database graph structure grows chains of equivalent states will emerge. These chains represent a mapping from a perceived state through an agent's available choices to a set of next states, the chain is simply a representation of how an agent class may influence its environment starting from a given state. This is illustrated in figure 6.19. Coaching agents will be able to infer the most influential

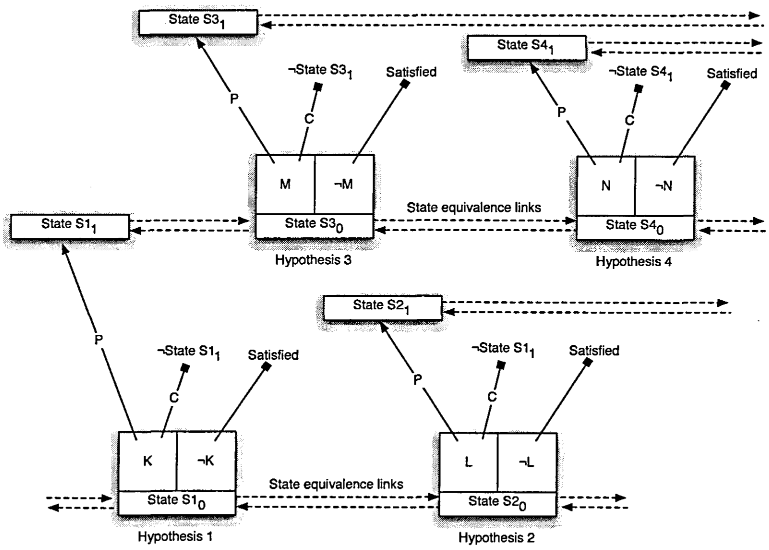


Figure 6.17: Coach database, illustration of possible multiple level hypotheses

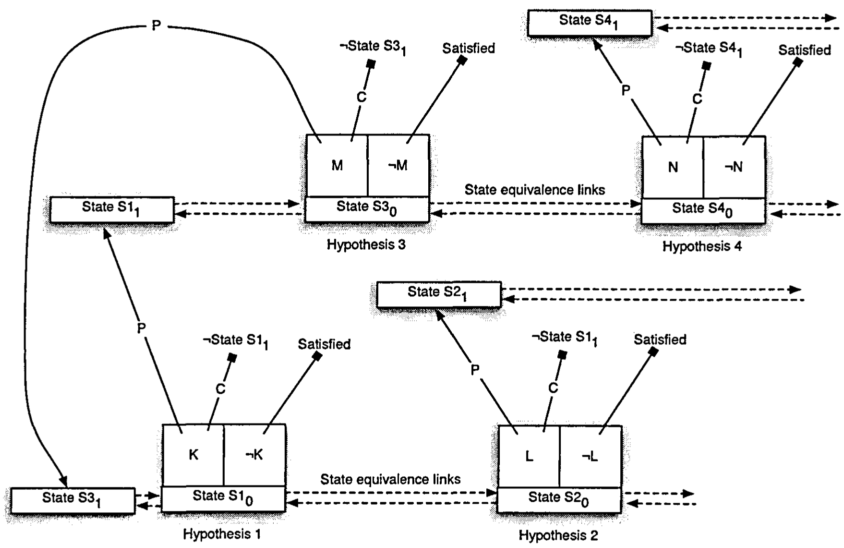


Figure 6.18: Coach database, admission of loops to extant states

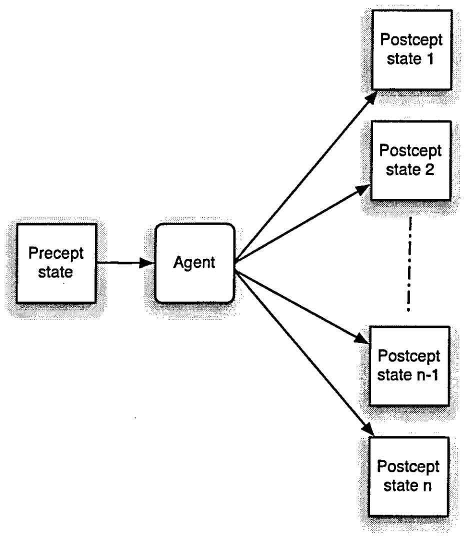


Figure 6.19: State equivalence chain as an agent internal mapping

action from a given state by walking state equivalence chains comparing the strength of each hypothesis by its positive and counter evidence tallies. These strengths will reflect the relative influence levels observed for each agent choice from the given starting state. In order to maximise single agent influence coaching agents will generate behaviour patches with weightings which bias agents towards choices that have been observed as having the most influence. We have mentioned a deontic thread in this work but can we say that simply biasing an agent towards a particular action has a deontic content? Horty [66, page 36] notes that standard deontic logic partitions future worlds into sets of ideal and non-ideal worlds yet we are biasing rather than requiring behaviour. Hansson [57] describes a *situationist* deontic logic which can take things as given, allowing an agent to make the best choice depending on its current circumstances. Girle [49, page 175] notes that Raymond Bradley (see Bradley and Swartz [17]) has suggested that the deontic O and P operators may be treated as qualified necessity and possibility operators and that the qualification may be moral or, in this case, practical. Anderson [4] considers OA, meaning that A is obligatory, may be reduced to $O(\neg A \supset V)$ where V indicates either that a norm has been violated and that the agent faces possible sanction or, in our case, reduced utility. Coaching agents will use the $O(\neg A \supset V)$ axiom which is a possible reading of obligation, they are not concerned with the mechanics of bringing a situation about but are, on the basis of observed evidence requiring that, agents act in a particular way.

6.6.5 Detecting evidence of serial influence, experimenting with the $P:C$ ratio

The coaching agent must generate new behaviour which it can seed in the environment. We have some guidelines, from section 5.1, which allow us to evaluate hypothesis strength. We assume that the negative evidence component is an enabling switch, if a hypothesis has no negative evidence then it is not fully formed and will not be used as a seed for new behaviours. We proceed to explore how this measure and switch may be used to rank hypotheses on their relative strength.

The investigation was by simple experiments and the first of these set three agents and three boolean variables or propositions in a single location environment. Each of the agents has four available actions, only one of these actions will change an associated proposition and the actions have a flat probability distribution. Agents operate only once per simulation cycle and in a random sequence. Noise is introduced into the cycle with a 10% chance of each proposition *flipping* its sense at a random point in the agent action sequence. Each proposition has a prior probability of 0.1 of being set true at the beginning of the simulation. Simulation cycles are non episodic in that the proposition values at the end of one cycle will be carried over to the beginning of the following cycle.

Noise will interfere with the results of agent actions but our privileged observer status affords us knowledge of which actions are genuinely influential. We know that $\gamma/M \rightsquigarrow Z$ and $\beta/L \rightsquigarrow Y$ are good hypotheses. α*'s* behaviour is a little more complex, its action K copies the value of Y onto X so that α has influence over X but this is contingent on β*'s* prior action. Our privileged world view tells us that $\beta/L;\alpha/K \rightsquigarrow X$ is a good hypothesis. Each agent has a flat behaviour probability function, each of the agent's four actions are equally likely to be selected. The 10% random flip noise each cycle is very significant compared with the 25% appropriate agent action chance that the flat probability distribution. This is a very noisy environment. The experimental data of table 6.2 presents a set of simple hypotheses for the scenario described above after

Table 6.2: Serial influence investigation after 1000 cycles, no ranking

Hypothesis	Evidence			P:C ratio	$P-C$ value	Sound
	Positive-	Negative-	Counter-			
$\alpha/K \rightsquigarrow X$	174	✓	67	0.721992	107	?
$\alpha/K \rightsquigarrow Y$	186	✓	55	0.771784	131	✗
$\alpha/K \rightsquigarrow Z$	187	✓	54	0.775934	133	✗
$\alpha/\neg K \rightsquigarrow X$	481	✓	278	0.633729	203	✗
$\alpha/\neg K \rightsquigarrow Y$	590	✓	169	0.777339	421	✗
$\alpha/\neg K \rightsquigarrow Z$	586	✓	173	0.772069	413	✗
$\beta/L \rightsquigarrow X$	334	✓	124	0.729258	210	?
$\beta/L \rightsquigarrow Y$	220	✓	9	0.960699	211	✓
$\beta/L \rightsquigarrow Z$	0	✗	0	0	0	✗
$\beta/\neg L \rightsquigarrow X$	498	✓	273	0.645914	225	✗
$\beta/\neg L \rightsquigarrow Y$	556	✓	215	0.721141	341	✗
$\beta/\neg L \rightsquigarrow Z$	596	✓	175	0.773022	421	✗
$\gamma/M \rightsquigarrow X$	147	✓	101	0.592742	46	✗
$\gamma/M \rightsquigarrow Y$	185	✓	63	0.745968	122	✗
$\gamma/M \rightsquigarrow Z$	230	✓	18	0.927419	212	✓
$\gamma/\neg M \rightsquigarrow X$	508	✓	244	0.675532	264	✗
$\gamma/\neg M \rightsquigarrow Y$	591	✓	161	0.785904	430	✗
$\gamma/\neg M \rightsquigarrow Z$	543	✓	209	0.722074	334	✗

1000 simulation cycles. These hypotheses are grouped by agent and are not ranked. Hypotheses that are known to be good – from privileged observer knowledge of the code used in the simulation – are marked by a "✓" in the rightmost column, those that are nonsensical, that is actions and results that are not explicitly coded and are the result of noise or overlapping percepts, are marked with a "✗" and those related to serial influence with a "?". The single agent hypotheses that we know are good, $\beta/L \rightsquigarrow Y$ and $\gamma/M \rightsquigarrow Z$, both exhibit high $P\colon C$ ratio values. Note that in this simulation the $P\colon C$ ratio is bounded to 1 and, consequently, 1 is a high value.

Early in the system run nonsensical negative hypotheses gathered high $P\colon C$ values. For example, $\gamma/\neg M \rightsquigarrow Y$ has a $P\colon C$ ratio of 0.785904 which makes it a possibly interesting hypothesis. Recall that our hypotheses are based on the condensed binary choice partitioning of definition 37. This means that when γ executes $\neg M$ it is simply executing another of its choices. Given that agents have a finite choice set we may say that the negation of action M implies the execution of some other action. Thus the refraining from M hypothesis, $\gamma/\neg M \rightsquigarrow Y$, is redundant because it is addressed by some other hypotheses based on commission of an action. Note that an agent's choice partition may include a null action and that, as Halbwachs [50] indicates, such actions are standard practice in synchronous reactive systems. The $\gamma/\neg M \rightsquigarrow Y$ hypothesis is simply a summary of the complement of M and $Choice_{\beta}$.

This brings us to the difficult question of what is a good $P\colon C$ ratio? Eliminating the negative hypotheses, for the reasons discussed immediately above, and ranking the remaining hypotheses by their $P\colon C$ ratio gives the data shown in table 6.3.

Table 6.3: Serial influence data after 1000 cycles, ranked by $P\colon C$ ratio

Hypothesis	Evidence			P:C ratio	$P - C$ value	Sound
	Positive-	Negative-	Counter-			
$\beta/L \rightsquigarrow Y$	220	✓	9	0.960699	211	✓
$\gamma/M \rightsquigarrow Z$	230	✓	18	0.927419	212	✓
$\alpha/K \rightsquigarrow Z$	187	✓	54	0.775934	133	×
$\alpha/K \rightsquigarrow Y$	186	✓	55	0.771784	131	×
$\gamma/M \rightsquigarrow Y$	185	✓	63	0.745968	122	×
$\beta/L \rightsquigarrow X$	334	✓	124	0.729258	210	?
$\alpha/K \rightsquigarrow X$	174	✓	67	0.721992	107	?
$\gamma/M \rightsquigarrow X$	147	✓	101	0.592742	46	×
$\beta/L \rightsquigarrow Z$	0	×	0	0	0	×

By inspection, the single agent influence hypotheses that we know are good have both risen to the top of the table and, even in this very noisy environment, are grouped closely together. There is a larger gap between the top two hypotheses and the next five which exhibit a closely grouped $P\colon C$ ratio. This is followed by another gap and two poorly scoring hypotheses. We are unable to state absolute $P\colon C$ ratio values but from table 6.3 we may infer some simple heuristics to guide a coaching agent in selecting good hypotheses. After a number of cycles, we assume that the top scoring hypothesis is good. From there we work down the table looking for relatively large gaps in the score and use these as a guideline for partitioning the table.

This relative gap heuristic allows us to identify candidates for seeding as single agent hypotheses, but what of the remaining hypotheses? From our privileged position we know that there are two related hypotheses which contribute to the serial influence that brings about X. α's reliance on an earlier action by β introduces additional noise – over and above that already in the system – to α's influence. This, as one would intuitively expect, causes the $P\!:C$ ratio for $\alpha/K \rightsquigarrow Y$ to be rather low. Note that the gap between the $P\!:C$ ratio for $\alpha/K \rightsquigarrow Y$ and the hypothesis below it in table is relatively large compared with the group of hypotheses above it and below those that we already believe to be good. This allows us to infer another relative gap heuristic which partitions the table into three blocks, the top block contains candidates for seeding as simple, single agent behaviours, the middle block contains hypotheses which warrant further investigation and the bottom block contains hypotheses that are uninteresting.

This partitioning of the table into three blocks defined by relative $P\!:C$ ratio gaps offers little, beyond the identification of possible components of other agent influences, to guide the coach in selecting behaviours to seed. The $P\!:C$ ratio offers one means of assessing noise in agent behaviours, we may use the same positive and counter evidence tallies to measure the absolute difference between the parameters.

Using the two heuristics outlined above to remove single agent candidate hypotheses and uninteresting hypotheses then ordering on $P-C$ value gives us table 6.4.

Table 6.4: Potential serial influence candidates after 1000 cycles, ranked by $P-C$ value

Hypothesis	Evidence			P:C ratio	$P-C$ value	Sound
	Positive-	Negative-	Counter-			
$\alpha/K \rightsquigarrow X$	174	✓	67	0.721992	107	?
$\gamma/M \rightsquigarrow Y$	185	✓	63	0.745968	122	×
$\alpha/K \rightsquigarrow Y$	186	✓	55	0.771784	131	×
$\alpha/K \rightsquigarrow Z$	187	✓	54	0.775934	133	×
$\beta/L \rightsquigarrow X$	334	✓	124	0.729258	210	?

The hypotheses that we know form part of the sequential influence chain lie at the top and bottom of the hypothesis table. $\alpha/K \rightsquigarrow X$ is the *delivery* action for the influence. If we naively coach α by biasing its chance of choosing action K to 35% (from the 25% of the flat probability distribution) we get the data of table 6.5.

Table 6.5: Potential serial influence with coached α (at 35%) after 1000 cycles, ranked by $P - C$ value

Hypothesis	Evidence			P:C ratio	$P - C$ value	Sound
	Positive-	Negative-	Counter-			
$\gamma/M \rightsquigarrow Y$	185	✓	63	0.745968	122	×
$\alpha/K \rightsquigarrow X$	248	✓	95	0.723032	153	?
$\alpha/K \rightsquigarrow Y$	264	✓	79	0.769679	185	×
$\alpha/K \rightsquigarrow Z$	265	✓	78	0.772595	187	×
$\beta/L \rightsquigarrow X$	332	✓	126	0.724891	206	?

Despite coaching α*'s* behaviour does not show any significant change in influence and, intuitively, this is to be expected. β*'s* earlier enabling action brings about a world state where α may more reliably exercise its influence and, consequently, β*'s* action is the more influential of the pair. Although α*'s* influence chain terminating action lies near the top of the $P - C$ ordered hypothesis block – indicating potential membership of an influence chain – β*'s* more influential enabling action lies at the bottom of the block. How, then, do we proceed?

6.6.6 Action preconditions, precepts and more focused hypotheses

So far, given a complete set of single agent/action hypotheses we have a heuristic which allows the identification of candidates for single agent influence and identifies uninteresting hypotheses. This leaves us with a set of potential candidates for serial influence behaviour but without a means for identifying the most interesting members of this set.

The hypotheses considered above were *global hypotheses* based on an agent's influence without regard to the current state of its world. Such hypotheses are blunt instruments which may capture the fact that an agent is influential but may not capture *when* that agent has influence in the world. If a coach detects that α/K has influence and seeds behaviours that have α executing K without regard to its current environment then the coach is seeding behaviours that potentially reduce agent influence because action K is only influential in certain circumstances and a blind bias towards K may be inappropriate.

To do this we need to refine our approach to evaluating hypotheses and move away from the rather blunt global approach outlined above. The actor agent operation cycle, described in definition 25, involves the agent gathering a set of precepts, executing some action from its available set of actions then generating a set of postcepts. All of these data are contained in history patches which provide the coaching agent's window on the world. The availability of precepts gives us a means of identifying more accurately where particular actions have influence.

The global hypothesis approach simply stated that α/K has influence over X regardless of α's precepts. We refine this by adding preconditions derived from agent precepts. The addition of these preconditions will increase the number of hypotheses that we have to manage but at the same time provides data about world state prior to an action. The experimental world is simple, having only three propositions, X, Y and Z (we neglect the presence of agents as this is a single location world). Considering the simple, single agent influence case of γ we have a set of global hypotheses, a set of hypotheses with X as a precept, a set of hypotheses with Y as a precept and a set of hypotheses with Z as a precept. These sets may, of course, overlap as it is quite possible for, say, both X and Z to be true in the agent precepts.

We start by considering the hypotheses relating to this simple influence. The data are collated in table 6.6 where hypothesis prefixes read GH for a global hypothesis, $X1$ where X is true in the precepts, $X0$ where X is false in the precepts and similarly for the other propositions Y and Z. This formidable table may be immediately trimmed by applying the heuristics outlined above and since we are only, at this point, interested in simple, single agent hypotheses we retain only the top set of hypotheses in each precept group. This filters out the noise driven propositions that γ has no influence over and is in line with our privileged knowledge of the system.

This leaves us with table 6.7 where we see that the $P\!:\!C$ ratio values are all very similar. This may be interpreted in two ways, looking at the precept grouped hypotheses the similarity between the $\top$ and $\bot$ values for X, Y and Z indicate that the values of these precepts have no significant effect on γ's ability

Table 6.6: $\gamma/M \rightsquigarrow Z$, complete set of global and precept prefixed hypotheses

Hypothesis	Evidence			P:C ratio	$P-C$ value	Sound
	Positive-	Negative-	Counter-			
Global hypotheses						
$GH\gamma/M \rightsquigarrow X$	151	✓	75	0.668142	76	×
$GH\gamma/M \rightsquigarrow Y$	182	✓	44	0.80531	138	×
$GH\gamma/M \rightsquigarrow Z$	212	✓	14	0.938053	198	✓
X precept hypotheses						
$X1\gamma/M \rightsquigarrow X$	131	✓	14	0.903448	117	×
$X1\gamma/M \rightsquigarrow Y$	123	✓	22	0.848276	101	×
$X1\gamma/M \rightsquigarrow Z$	136	✓	9	0.937931	127	✓
$X0\gamma/M \rightsquigarrow X$	20	✓	61	0.246914	-41	×
$X0\gamma/M \rightsquigarrow Y$	59	✓	22	0.728395	37	×
$X0\gamma/M \rightsquigarrow Z$	76	✓	5	0.938272	71	✓
Y precept hypotheses						
$Y1\gamma/M \rightsquigarrow X$	136	✓	41	0.768362	95	×
$Y1\gamma/M \rightsquigarrow Y$	164	✓	13	0.926554	151	×
$Y1\gamma/M \rightsquigarrow Z$	164	✓	13	0.926554	151	✓
$Y0\gamma/M \rightsquigarrow X$	15	✓	34	0.306122	-19	×
$Y0\gamma/M \rightsquigarrow Y$	18	✓	31	0.367347	-13	×
$Y0\gamma/M \rightsquigarrow Z$	48	✓	1	0.979592	47	✓
Z precept hypotheses						
$Z1\gamma/M \rightsquigarrow X$	121	✓	56	0.683616	65	×
$Z1\gamma/M \rightsquigarrow Y$	145	✓	32	0.819209	113	×
$Z1\gamma/M \rightsquigarrow Z$	167	✓	10	0.943503	157	✓
$Z0\gamma/M \rightsquigarrow X$	30	✓	19	0.612245	11	×
$Z0\gamma/M \rightsquigarrow Y$	37	✓	12	0.755102	25	×
$Z0\gamma/M \rightsquigarrow Z$	45	✓	4	0.918367	41	✓

to influence Z. An alternative reading is that the given closeness in value of the $P{:}C$ ratio for the global hypotheses we apply Occam's razor and take the simplest option - the global hypothesis. Either of these leads to the same conclusion that the precepts don't have any effect on γ's ability to bring about Z.

Table 6.7: $\gamma/M \rightsquigarrow Z$, minimal set of global and precept prefixed hypotheses

Hypothesis	Evidence			P:C ratio	$P - C$ value	Sound
	Positive-	Negative-	Counter-			
Global hypotheses						
$GH\gamma/M \rightsquigarrow Z$	212	✓	14	0.938053	198	✓
X precept hypotheses						
$X1\gamma/M \rightsquigarrow Z$	136	✓	9	0.937931	127	✓
$X0\gamma/M \rightsquigarrow Z$	76	✓	5	0.938272	71	✓
Y precept hypotheses						
$Y1\gamma/M \rightsquigarrow Z$	164	✓	13	0.926554	151	✓
$Y0\gamma/M \rightsquigarrow Z$	48	✓	1	0.979592	47	✓
Z precept hypotheses						
$Z1\gamma/M \rightsquigarrow Z$	167	✓	10	0.943503	157	✓
$Z0\gamma/M \rightsquigarrow Z$	45	✓	4	0.918367	41	✓

6.6.7 Using precept prefixed hypotheses to detect serial influence

What, then, of the serial chain which we know is good $-\beta/L;\alpha/K \rightsquigarrow X$? Collating the evidence for the delivery action gives us the data of table 6.8. As above this is a formidable table so we begin by attempting to trim it to a more amenable size.

Recall that the $P:C$ ratios for α's influence over X, Y and Z fell into the second group after applying our heuristics above. Here we see that the global hypotheses all have similar $P:C$ ratios and we move to the $P - C$ value, $GH\alpha/K \rightsquigarrow X$ has the lowest value indicating that it is the most interesting of the group and that this hypothesis may be contingent on something else. The coaching agent has no knowledge of the physics of its world and it may only infer that α's setting of X is dependent on β's earlier setting of Y from observation.

There are two ways that we may approach this, we may either select the most interesting global hypothesis and use that as a filter on the precept prefixed hypotheses or we may treat the global hypotheses as equals and rank all of the precept prefixed hypotheses.

Table 6.8: Influence delivery, $\alpha/K \leadsto X$, complete set of global and precept prefixed hypotheses

Hypothesis	Evidence			P:C ratio	$P-C$ value	Sound
	Positive-	Negative-	Counter-			
Global hypotheses						
$GH\alpha/K \leadsto X$	178	✓	63	0.738589	115	?
$GH\alpha/K \leadsto Y$	188	✓	53	0.780083	135	×
$GH\alpha/K \leadsto Z$	191	✓	50	0.792531	141	×
X precept hypotheses						
$X1\alpha/K \leadsto X$	121	✓	30	0.801324	91	?
$X1\alpha/K \leadsto Y$	126	✓	25	0.834437	101	×
$X1\alpha/K \leadsto Z$	118	✓	33	0.781457	85	×
$X0\alpha/K \leadsto X$	57	✓	33	0.633333	24	?
$X0\alpha/K \leadsto Y$	62	✓	28	0.688889	34	×
$X0\alpha/K \leadsto Z$	73	✓	17	0.811111	56	×
Y precept hypotheses						
$Y1\alpha/K \leadsto X$	165	✓	20	0.891892	145	?
$Y1\alpha/K \leadsto Y$	167	✓	18	0.902703	149	×
$Y1\alpha/K \leadsto Z$	145	✓	40	0.783784	105	×
$Y0\alpha/K \leadsto X$	13	✓	43	0.232143	-30	?
$Y0\alpha/K \leadsto Y$	21	✓	35	0.375	-14	×
$Y0\alpha/K \leadsto Z$	46	✓	10	0.821429	36	×
Z precept hypotheses						
$Z1\alpha/K \leadsto X$	143	✓	47	0.752632	96	?
$Z1\alpha/K \leadsto Y$	147	✓	43	0.773684	104	×
$Z1\alpha/K \leadsto Z$	177	✓	13	0.931579	164	×
$Z0\alpha/K \leadsto X$	35	✓	16	0.686275	19	?
$Z0\alpha/K \leadsto Y$	41	✓	10	0.803922	31	×
$Z0\alpha/K \leadsto Z$	14	✓	37	0.27451	-23	×

6.6.8 Filtering prefixed hypotheses against one global hypothesis

In the first case we select $GH\alpha/K \leadsto X$ as a target because it has the lowest $P-C$ value. and extract all of the precept prefixed hypotheses for $\leadsto X$. This gives the data of table 6.9, ordering these by $P:C$ ratio places $Y1\alpha/K \leadsto X$ at the top of the table.

Table 6.9: Influence delivery, precept prefixed hypotheses filtered on single global hypothesis

Hypothesis	Evidence			P:C ratio	$P-C$ value	Sound
	Positive-	Negative-	Counter-			
Global hypothesis						
$GH\alpha/K \rightsquigarrow X$	178	✓	63	0.738589	115	?
Ordered $\rightsquigarrow X$ hypotheses						
$Y1\alpha/K \rightsquigarrow X$	165	✓	20	0.891892	145	?
$X1\alpha/K \rightsquigarrow X$	121	✓	30	0.801324	91	?
$Z1\alpha/K \rightsquigarrow X$	143	✓	47	0.752632	96	?
$Z0\alpha/K \rightsquigarrow X$	35	✓	16	0.686275	19	?
$X0\alpha/K \rightsquigarrow X$	57	✓	33	0.633333	24	?
$Y0\alpha/K \rightsquigarrow X$	13	✓	43	0.232143	-30	?

There is a relatively large $P:C$ ratio gap between the first hypothesis and the others in the table. This indicates that having proposition Y holding before α/K holds the most promise. The coaching agent already has data, in table 6.3, indicating that $\beta/L \rightsquigarrow Y$ is a good hypothesis which leads to the Y that the coach infers is a required precondition for α being able to bring about X. The coach will then generate a behaviour patch which increases β's bias towards selecting action L.

6.6.9 Selecting behaviour from ordered prefixed hypotheses

The second approach is to take the set of global hypotheses as is, rank the prefixed hypotheses in table 6.8 by $P:C$ ratio and select the most promising from the ranked group. These data are displayed in table 6.10.

Using the relative difference heuristic we see the first large gap between hypotheses three and four. Within this top group we see the effects of the system being very noisy. We know, because of our privileged knowledge, that the two top ranked hypotheses do not hold but the noisiness of the environment provides evidence that they do. The good hypothesis, that with Y holding as a precept, is in third place but still falls into the interesting group. In this case the coach would seed three behaviours, one leading to Z, one leading to X and one leading to Y.

Table 6.10: Influence delivery, precept prefixed hypotheses ranked by $P:C$ ratio

Hypothesis	Evidence			P:C ratio	$P-C$ value	Sound
	Positive-	Negative-	Counter-			
$Z1\alpha/K \leadsto Z$	177	✓	13	0.931579	164	×
$Y1\alpha/K \leadsto Y$	167	✓	18	0.902703	149	×
$Y1\alpha/K \leadsto X$	165	✓	20	0.891892	145	?
$X1\alpha/K \leadsto Y$	126	✓	25	0.834437	101	×
$Y0\alpha/K \leadsto Z$	46	✓	10	0.821429	36	×
$X0\alpha/K \leadsto Z$	73	✓	17	0.811111	56	×
$Z0\alpha/K \leadsto Y$	41	✓	10	0.803922	31	×
$X1\alpha/K \leadsto X$	121	✓	30	0.801324	91	?
$Y1\alpha/K \leadsto Z$	145	✓	40	0.783784	105	×
$X1\alpha/K \leadsto Z$	118	✓	33	0.781457	85	×
$Z1\alpha/K \leadsto Y$	147	✓	43	0.773684	104	×
$Z1\alpha/K \leadsto X$	143	✓	47	0.752632	96	?
$X0\alpha/K \leadsto X$	57	✓	33	0.633333	24	?
$X0\alpha/K \leadsto Y$	62	✓	28	0.688889	34	×
$Z0\alpha/K \leadsto X$	35	✓	16	0.686275	19	?
$Y0\alpha/K \leadsto Y$	21	✓	35	0.375	-14	×
$Z0\alpha/K \leadsto Z$	14	✓	37	0.27451	-23	×
$Y0\alpha/K \leadsto X$	13	✓	43	0.232143	-30	?

6.7 What does a coaching agent do?

In this chapter we have examined the relationship between a coaching agent and the actor agents in its environment. Coaching agents will aggregate observations by collecting partial history traces from actor agents. These traces are assembled into a database structure which links hypotheses as a directed cyclic graph and attaches evidence tallies to each observed transition. The coaching agent will use simple heuristics to partition the set of database hypotheses into potentially influential hypotheses – which are used as a basis for synthesising new agent behaviours – and uninteresting hypotheses which are recorded and updated just in case they become interesting at some point in the future.

Chapter 7

Exploring influence, implementation and experiments

We have developed a theory of influence, described *leadsto* and *may lead to* operators and built a partial logical characterisation of them. This characterisation has two purposes; to explore the behaviour of influence and to guide us in implementing coaching agents within a system. The previous chapter explored some of the theoretical machinery required by a coaching agent to allow it to interpret observations as meaningful indicators of agent influence. This chapter moves the theoretical machinery in to a practical system and we describe an experimental system. We are concerned, for the most part, with theoretical aspects of influence and do not cover implementation aspects in great detail although, where necessary, we discuss architectural points so as to outline how the agents operate and to indicate that they follow our theory very closely.

7.1 Experimenting with influence, building bridges

One of the main objectives for grounding this work in the computational domain was allow us to carry out experiments using our theory of influence. Our intentions are threefold, we intend to demonstrate that coaching agents are able to synthesise behaviours from their operations. We intend to show that different agent classes or types may operate in such a way as to jointly bring about something that individual classes are unable to achieve and that this may be represented in a directed hypothesis graph. We wish, also, to demonstrate that agents may use influence additively in order that a number of agents acting in series or parallel may bring about something that individuals are unable to do.

A series of experiments was constructed to test these possibilities and bridge building provided a setting where influence, an agent partially bridging a gap, may be extended to completely bridge a gap. The most attractive aspect of this is that agents need not necessarily see that they are aiming for a target, a gradually bridged gap is seen as gradually extending influence until it allows agents to reach a new part of their world.

Building a virtual bridge is a formidable undertaking and since we are not interested in the physics of bridge building we reduce the concept to that of a numerical accumulator. This allows us to approach a simulation of bridge building without the burden of replicating real world physics and without losing any of the constraints and limitations imposed by that real world. Consider a world with agents which may either add a unit to or subtract a unit from some form of accumulator. Such units may be thought of as representing components of a bridge extending to span a gap between lands.

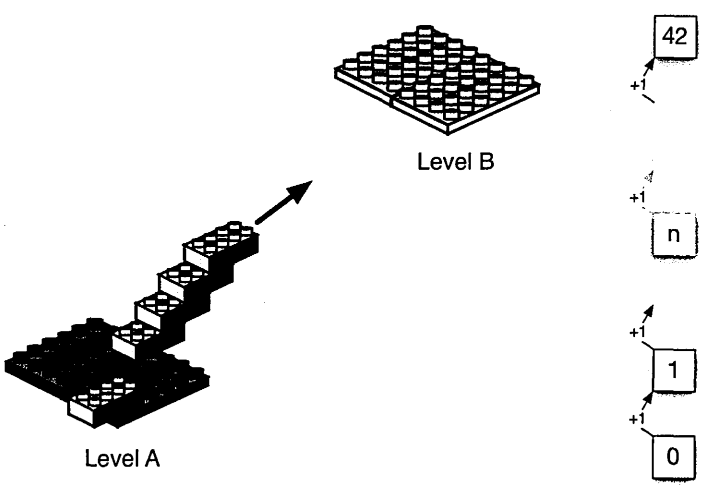


Figure 7.1: Physical and numerical bridge building

The notion of normative systems has been mentioned frequently in earlier parts of this work and here this notion becomes concrete in a simple system. If the definition for this world states that, for example, 42 is good then this will be the state that an observer sees as an indication that agents and the system have achieved something. The coach has no notion of 42 and operates only by maximising observed influence. When a coach sees a transition from 0 to 1 it sees a change from 0 to something and that represents influence. We can see that norms are system dependent and in complex systems they may also be location dependent with that location possibly being a virtual location. The observer agent "knows" system norms but coaching agents do not, coaching agents are only aware that incrementing an accumulator is an influential action, the observer sees this influential action leading to a system norm being satisfied. This observer / coach relationship implicitly captures the ought implies can deontic identity. Similarly in a physical bridging environment an observer may know that an agent being on level B is good and a coach may view a new agent percept as evidence of some influence. A coach may have no notion of the spatial relationship between level A and level B or that it is possible for agents to bridge the gap and get from one to the other.

We noted in section 2.12 that we make no claims for methods of identifying or predicting emergent behaviour and this is clear from these bridging examples. The only tool that a coaching agent has is its observation of influence and it uses this as a constraining device on the state space search.

Bridges and accumulators differ in that the physical bridge has a spatial element. A bridge may be the result of an agent exercising influence but if the bridge is not in a good location then the bridge may not open access to a previously inaccessible part of the world. Figure 7.2, for example, shows two bridges, bridge B1 will eventually reach level B but the green agents are going in a direction which will not lead to a new part of the world.

In such cases we expect that by inspecting the directed graph structure of its database a coaching agent will, eventually, see that chains of events have necessary preconditions. In figure 7.2 we see that the agents building bridge B1 have started their bridge building at the edge of a chasm and this will lead to level B. The agents building bridge B2 have started their bridge fully on land and although in some cases this may

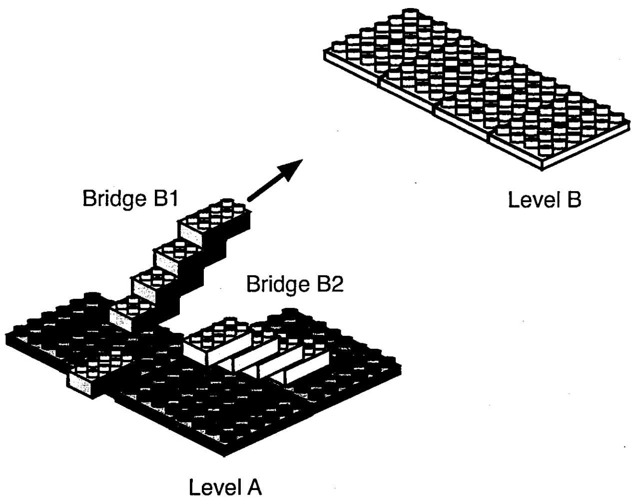


Figure 7.2: Bridges may be built in the wrong direction

bridge a gap there will be cases where it does not and this will be reflected in evidence tallies on the directed graph.

The arithmetic accumulator may model a single, simple bridge and show influence being used constructively to extend a bridge span by increasing the value of an accumulator. This is sufficient for a single cell environment where the good state may be a bridge of a given length or accumulator of a certain value. In a multi cell environment we need some means of simulating effective and ineffective bridges. This may be achieved by using a variety of classes of accumulator distributed throughout the environment and, because this is a multi cell environment, agents may move between cells. Agents will be able to identify accumulator classes – a red accumulator or a green accumulator – and certain of these classes will be bounded. Some may be bounded below the observer's target of 42 and some above. In the physical environment we would expect the system to learn not to build bridges facing inland and in the accumulator environment we expect agents to learn which accumulators are bounded and will not lead to the good state.

These scenarios are attractive for a number of reasons. The observer's system goal states are easily defined and easily observed. The coaching agent's lack of fundamental knowledge means that it is unable to see that agent behaviour is approaching a solution until the observer sees that it has been achieved.

7.2 A series of experiments: looking for extended influence

Our experiments in bridge building were carried out over a series of experiments. These ranged from the almost trivially simple to very complex. By complexity we mean a combination of the size of the state space that contains a solution and number of pathways that an agent may follow through that state space. We start with a single cell, single agent world. We use the same agents throughout the experiments so actions that would cause the agent to move to another location are cast as null actions in single cell environments. One of the coaching agent's abilities that we are keen to test is its ability to detect influence in noisy environments. We do this by running each experiment in four different configurations, with and without noise, and with and without coaching. Each of these runs was repeated a number of times using different random number generator seeds so as to generate a spread of results and reduce the chance of our happening to simply get a good sequence from the randomiser. Early experimental work revealed quirks of random number generator behaviour in a multithreaded Windows environment so our experiments use a single instance of a random number generator based on an algorithm described by Leva [79] to guarantee a single sequence rather than a number of the same sequences.

The noise free, uncoached runs are intended to test stochastic ability, that is the possibility of agents finding a solution behaviour by randomly walking the state space. If the agents do find a solution then this will give a metric for gauging the improvement that coaching brings. The noisy uncoached runs are intended to test stochastic ability in the presence of random disturbances. The noisy and noise free coached runs are intended to test our theories that influence may be used to guide state space searching and, in comparison with the uncoached runs, to give an idea of the degree of improvement in search effort that coaching brings.

7.3 What do we see when we see influence?

What do we treat as influence in this experiment? Here we must be cautious so as to keep preconceptions and assumptions related with numerical values out of the coaching operation. This is a bridge building exercise and not a numerical state space exploration so the coaching agents neglect the value of the accumulator.

Recall that it is an external observer that sees good states and not the coach so the coaching agent's lack of knowledge of absolute accumulator value is no hindrance.

This classification of influence may be viewed as *a priori steering* towards a good result and this would certainly be the case if the aim of the experiment was to count up to a certain value. We have mentioned, above, that the transition from 0 to 1 is seen as influence and that this is because it brings about a new thing in the environment, something that has not been seen before. If we cast the accumulator changes as forward or backward moves along an under construction bridge, as in table 7.1, then this reading of accumulator value transitions makes sense. Stripped of numerical preconceptions and assumptions these readings are

Table 7.1: Influence in a bridge building world

Observation	Reading	Reason
Transition from zero	Influence	Forwards move to new state
Transition to zero	No influence	Backwards move from previous influence
Increased value	Influence	Forwards move to new state
Decreased value	No influence	Backwards move from previous influence

obviously not an attempt to give the system an *a priori* steer towards a desired result. A decrementing move will take the system back to an old state, something that has already been seen by the coach and is, consequently, a non influential action.

The 42 question or the question of an observed good state is a little more involved. In figures 7.1 and 7.2 we represented bridges as extending from level A to level B. If agents were able to perceive and report the type of cell that they were occupying then a coaching agent would be able to detect that bridge acts as a conduit to increased influence. Before the bridge was constructed agents had not perceived level B cells so the bridge has enabled agent exploration of a part of the environment, part of their state space, that was previously unreachable. This is an interesting case where agent influence does not necessarily change the environment – there always were two levels – but increases the size of the state space potentially making new good states accessible. This is an instance of the *gateways* that we illustrated in chapter 3, figure 3.2 (which is reproduced in this chapter as figure 7.4), aggregate agent influence opens gateways into previously inaccessible behaviour domains. This is something that will be explored in further work (briefly described in section 8.4). In a physical bridge building world there may be an implicit bounding to a bridge length. When agents are on a bridge their choices are limited, they may move forwards, move backwards or build forwards. If an agent maintains a full set of choices – including move left and move right – then it may fall off of the bridge and out of the experiment. This does not present difficulties, agents that fall off of bridges are lost to the system so a coaching agent will not see evidence of *move left* actions whilst on a bridge. This lack of evidence means that the coaching agent will not seed behaviour biasing agents towards falling off of bridges.

Our early work on influence indicated that other agent influence was observable by changes in agent choice partitioning. Logie et al. [83] described the effect of other agent action on agent choice partitioning by saying that influence was detectable where there was, when viewed in a branching time frame, a history available that was not available without other agent action. This was observable in parallel action cases, equations 3.1 and 3.2, and in the serial case of equation 3.3 (reproduced here as equations 7.1, 7.2, and 7.3).

$$\exists h.Choice^{m}_{\alpha\|\beta}(h) \subsetneq Choice^{m}_{\alpha}(h) \tag{7.1}$$

$$\exists h.Choice^{m}_{\alpha\|\beta}(h) \subsetneq Choice^{m}_{\beta}(h) \tag{7.2}$$

$$\exists h.Choice^{m}_{\gamma;\delta}(h) \subsetneq Choice^{m}_{\delta}(h) \tag{7.3}$$

In the practical application domain these equations presuppose some degree of awareness of choice partitioning somewhere in the system. Our actor agents are extremely simple with no awareness of choice partitioning in their make up. Indeed, their choice partitions are fixed in our experimental system of purely reactive agents. How, then, do we implement them in such a way as to be amenable to coaching? We now briefly consider some aspects of the implementation of coaching and actor agents before considering experimental details more fully.

7.4 Coaching agents - some implementation details

In chapter 6 we examined theoretical aspects of coaching agent behaviour. Here we briefly outline some practical aspects. The coaching agent's main task is to randomly sample history data in the environment, use that data to generate or update hypotheses which are used to generate and seed new behaviours in the environment. Hypotheses are ranked by two criteria, the $P\colon C$ ratio and the absolute $P - C$ value. The former is a strong indication of single agent ability and the latter is an indication of potential other agent nested abilities. Coaching agents do not have abilities to manipulate physical objects in the environment but can work with a special class of objects – data and behaviour patches – which have no effect in the environment and are used as a simple locale based communications channel between coaching agents and actor agents.

When a coach picks up an agent history patch it examines it for evidence of change and we consider agent driven change as prima facie evidence of influence. If the history patch indicates change then the coach will use it to either generate a hypothesis or, if it already has a matching hypothesis, update evidence data. History patches provide two types of evidence, evidence directly related to a hypothesis and what we term collateral evidence where data may be used to toggle the negative condition of non related hypotheses. For example, if a patch indicates that $\alpha/K \rightsquigarrow A$ then related hypotheses will be those where α/K holds so that the patch will provide positive evidence for a hypothesis that $\alpha/K \rightsquigarrow A$ and will provide counter evidence for other α/K hypotheses, for example $\alpha/K \rightsquigarrow B$. If a matching hypothesis is not present then the coach generates and stores a new hypothesis. Note that a single history patch may generate multiple hypotheses if there is evidence of multiple changes in the precept and postcept data.

After dealing with related hypotheses the coach will then use the data as collateral evidence to update counter or negative evidence for each applicable hypothesis in the database.

Intuitively the notion of coaching is that a coach sees that a particular action or set of actions is good in certain circumstances and generates an agent choice set which is biased towards those actions. Note that this is a bias and not a rule that in given circumstances an agent must execute a particular choice. Where the coach believes that certain choices will bring about a bad state then it may disable that choice by giving it a probability of 0. This represents our earlier intuitive description of Lamport's [78] safety and liveness properties which stated that something bad will not happen and that something good will eventually happen.

7.5 Actor agents - some implementation details

Recall that our fundamental notion of an agent is as a partitioning mechanism overlaid on a discrete branching time frame and that influence is where agent choice partitions the future. Recall, also, that our actor agents are purely reactive and have no notion of their abilities or choice partitioning. We intend to have coaching agents synthesise new behaviours for these simple, reactive actors and have the actors use these behaviours appropriately. How can we make these three seemingly incompatible properties and requirements work together?

In section 6.6.7 we discussed percept prefixed hypotheses. The notion was that given an agent class and a particular set of precepts then a given action has been observed as the most influential. In chapter 1.1.5 we noted that agents had fixed abilities and were individually unable to acquire new behaviours. Combining these fixed abilities with precept prefixed hypotheses allows us to consider actor agents as a mapping of precepts on to preferred behaviour patterns. This may be represented by the agent architecture illustrated in figure 7.3 where an agent consists of a collection of behaviours – each having a preferred action – selected by some precept filtering mechanism. Conceptually actor agents are a collection of finite state machines each with a weighted stochastic transition selection mechanism.

The agent, illustrated in figure 7.3, has six choices available to it, $Choices = \{K, L, M, N, O, P\}$. Each of these choices represent some action by that agent, this may be an individual action, an action in concert with other agents or a null action. The set of choices remains constant throughout an agent's life, it cannot acquire new choices and it cannot discard any of its current set of choices. The agent's choice set is overlaid by a set of weightings with each weighting assigning a preference for a single choice. Choice set 0, in figure 7.3, gives each of the elements of *Choices* an equal weighting, this is the agent's default state and it has no individual behaviour characteristics, each of its actions has an equal chance of being chosen. When the actor

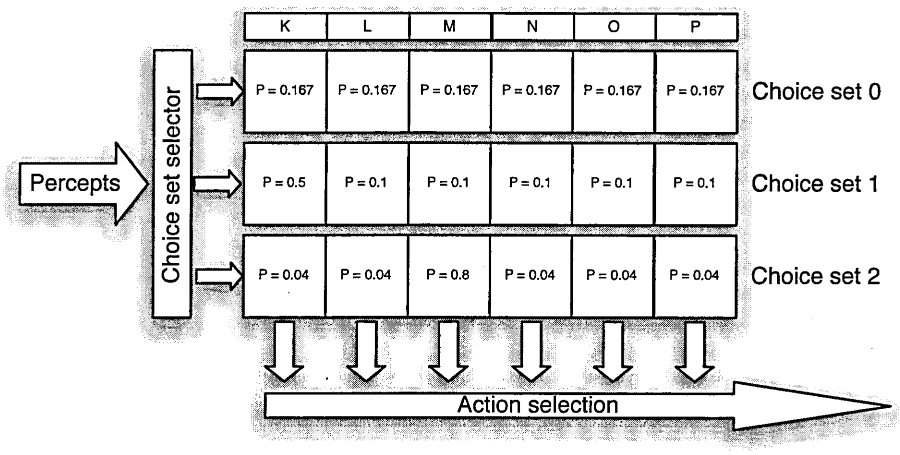


Figure 7.3: Agent internals - behaviour stack holding three behaviours

agent is operating it simply generates a random number, checks this against the weightings in the choice set and executes the appropriate choice.

If a coaching agent generates a new behaviour then this is packaged along with a precept template as a behaviour patch and dropped in the agent environment. The agent architecture of figure 7.3 has a single choice set attached to each percept slot. It would not make sense for an individual agent to hold multiple behaviours for a single precept trigger. We do expect that in complex worlds different classes of agents will emerge, classes that will select different actions for the same set of precepts and this is briefly discussed in section 8.4.2.

When an agent picks up a behaviour patch it uses the percept template as an identifying tag, if the agent has an older behaviour patch which uses the same tag then the newer choice weightings will replace the extant set. If the agent does not have a matching tag then it simply adds the new choice weighting set to its behaviour stack. For a given set of percepts an agent will select a behaviour template that has been generated by a coach and this will increase the likelihood that the actor will execute the coach selected action. We use the term biases quite deliberately, action selection is still a stochastic operation. In figure 7.3, choice set 1 has a probability of 0.5 for action K and choice set 2 has a probability of 0.8 for action

M. In both cases there is still a probability of the agent selecting an action from the complement set of the preferred action. This element of chance is deliberately included so as to cause some dither in the system preventing it from settling into locale specific behaviours – local minima and local maxima. Where there is no percept template match the agent simply selects a behaviour using a default behaviour template with a flat preference weighting.

This agent architecture is simply an extension of a standard perceive–select–act reactive agent. Instead of mapping percepts on to an action it maps them on to a biased action selection. By altering the bias settings for each behaviour set we are able to coach the agent and alter its behaviour without compromising its simple reactive structure or leading it along the path towards cognitive agency.

7.6 Bridge building, a single agent - single location simulation

A first experiment was carried out in a simple, single cell world If the agent's actions were fully deterministic then the simulation would be a trivial investigation of cause and effect. We introduce noise into the system in a similar manner to that of the noise of earlier experiments. There is a chance that noise may be introduced at a random point in the cycle and if it is it will increment or decrement the accumulator by a small amount and will not decrement it below zero. The agent's action cycle (neglecting behaviour and history patch management) is in three steps, precepts, action and postcepts and noise may occur at any gap in this cycle meaning that noise may occur before or after agent action. Our expectations at this point are limited, on uncoached runs we may see evidence of simple *stochastic ability*. On coached runs we expect to see the coaching agent detect influence, as described in table 7.1, and seed appropriate behaviours so as to encourage actors to exercise this influence.

7.6.1 Single agent, single cell results

The first step was to run a raw uncoached agent in a noise free environment to get an idea of its stochastic ability and the requirements for a random state space search. Being uncoached the actor agent's behaviour weightings remained flat throughout. The actor agent has seven actions which are defined, for the convenience of human observers, as:

```
#define NULL_ACTION_1   0
#define DECREMENT       1
#define NULL_ACTION_2   2
#define NULL_ACTION_3   3
#define NULL_ACTION_4   4
#define INCREMENT       5
#define NULL_ACTION_5   6
```

Because this is a single cell setting there are five null actions, when the experiment moves to a multi cell domain then four of these will be replaced by move actions leaving one null. If an uncoached agent were left running in a noiseless environment then one would expect any accumulator to, on average, stay at or close to zero. If the distribution of actions is relatively flat then the number of increment calls will be roughly equal to the number of decrement calls and the null actions may be neglected. This experiment was

Table 7.2: Results: single agent, single cell, no noise and no coaching

Cycles	Seed 1	Seed 2	Seed 3	Seed 4
100	3	2	4	2
1000	30	4	20	6
10000	65	13	39	27

carried out using four different seeds for the random number generator and the accumulator values tallied at 100, 1000 and 1000 cycles. The results, shown in table 7.2, are as expected. Even if the increment and decrement operations were perfectly balanced the fact that the accumulator will not decrement below zero should give a small advantage to the increment operation and the granularity of the choice (based on integer variables) means that the randomness of choice is not perfect. An average value of 36 after 10000 cycles is not unreasonable. Randomly behaving agents with a flat choice distribution operating in an environment that slightly favours incrementing will cause the value stored in an accumulator to rise at a very slow rate.

Noise was introduced and the procedure was repeated using the same random number seeds in each case. Note that even though the seeds are the same this does not mean that the agent's choice sequence is identical to that of the noise free setting. The same random number generator is now being given additional noise generating duties so the sequences generated by the same seeds are now spread across a number of tasks. The noise was introduced at a single randomly chosen point in the agent cycle. Instead of a clean *precepts, action, postcepts* cycle we now have *potential noise, precepts, potential noise, action, potential noise, postcepts, potential noise* cycle. In practice, noise is only introduced at one place in each cycle. At each noise insertion point there is a fifteen percent chance that the system will actually be noisy – this level was selected randomly and chosen to be close to the agent's chance of a particular choice using a flat weighting – once again this is a large noise component. Noise takes the form of a random shift of ± 3 units – this was, again, selected randomly – with the accumulator having a lower bound of zero. The data for

Table 7.3: Results: single agent, single cell, with noise and no coaching

Cycles	Seed 1	Seed 2	Seed 3	Seed 4
100	0	7	9	37
1000	57	11	49	34
10000	47	28	83	36

an uncoached agent in a noisy environment are shown in table 7.3. Noise appears to bring a degree of non linearity into the simulation. This is most, and surprisingly so, obvious for the fourth set of results. The average accumulator value after 10000 cycles is 48, greater than that of the noise free examples but the obvious non linear progression indicates that these values are more volatile and that detecting influence in such a setting will be more difficult.

Coaching was introduced on the basis of table 7.1 so that increasing accumulator values are seen as influential in that they are new states of the world.

Table 7.4: Results: single agent, single cell, noise free with coaching

Cycles	Seed 1	Seed 2	Seed 3	Seed 4
100	10	9	12	16
1000	110	109	141	93
10000	1037	999	1048	963

Coaching in these single cell settings is what we term *hands off coaching*. When a coaching agent synthesises a behaviour it does so once only, there is no continual monitoring and updating of a synthesised behaviour to strengthen it on repeated evidence of influence. *Hands on coaching*, described below, continually modifies behaviours giving a buffered positive feedback loop to prevent saturation by a single choice.

7.6.2 Examining what agents are doing

Looking beyond the tabular data to examine agent log messages in a single agent coached run. These indicate what agents see or do at various points in their operation cycle and by examining them we we see the steps involved in the synthesis and distribution of a new behaviour.

```
Agent Bingo cycle 2: No new behaviours.
Agent Bingo cycle 2: Precepts gathered.
Agent Bingo cycle 2: No coached behaviours, selecting default behaviour.
Agent Bingo cycle 2: Increment action.
```

The agent cycle begins with a check for new behaviours, there are none here so it continues to gather its precepts and will select an action uses its default flat behaviour weighting, this time it selects an *increment* action.

```
Accumulator 0 incremented, value is now: 1.
Accumulator 0 cycle 2 no noise.
Agent Bingo cycle 2: Postcepts gathered.
```

The accumulator is incremented and the agent gathers its postcepts. Note that the accumulator value is displayed here for convenience and that coaching agents use only transition data.

```
Agent Bingo cycle 2: History patch start --->
This patch tag = Bingo:2
Precepts -  Ac: yes.  AcClass: 1     zero.
Action = 5.
Postcepts -  Ac: yes.  AcClass: 1 not zero.
Agent Bingo cycle 2: History patch end <--
```

The history patch is now in the environment and may be collected by a coaching agent.

```
Coach (cycle 2): Patch showing influence picked up.
Accumulator 0 cycle 2 value = 1
```

The coaching agent collects the history patch and detects that it exhibits influence because of the increased accumulator value.

```
Patch generator - synthesised new behaviour (flat) #0.
Patch generator - positive tweak for behaviour 0.
Patch generator - dropped behaviour 0.
```

A new patch is synthesised and a *positive tweak* is applied to increase the weighting of the increment action. This behaviour is placed in the environment for an actor agent to collect.

```
Agent Bingo cycle 4: New behaviour added.
Agent Bingo cycle 4: Precepts gathered.
Accumulator 0 cycle 4 no noise.
Agent Bingo cycle 4: No applicable behaviours, selecting default behaviour.
```

The actor agent collects and installs the new behaviour at the start of the following cycle. This patch was generated on a *zero* to *not zero* transition so it's precept requirements are that an accumulator be present and that that accumulator be zero. In this case the accumulator is non zero, the agent's precepts do not match the patch trigger precepts so the agent defaults to a flat behaviour weighting.

```
Agent Bingo cycle 15: New behaviour added.
Agent Bingo cycle 15: Precepts gathered.
Agent Bingo cycle 15: Selecting coached behaviour #1.
Agent Bingo cycle 15: NULL action #2.
Agent Bingo cycle 15: Postcepts gathered.
```

Later in the simulation the coaching agent generates a new behaviour with a different set of precepts which indicate that it is suitable for a non zero accumulator. This is added to the agent's behaviour slot 1 (zero based) and this behaviour's weighting, which has a bias towards the increment action, is consistently selected in following cycles. In this *hands off coaching* setting the bias is simply an additional 10% chance of the increment action being selected. Despite this small additional weighting the results produced by the agent, shown in table 7.5, are significantly better than those of uncoached agents as shown in table 7.3.

Table 7.5: Results: single agent, single cell, noisy with coaching

Cycles	Seed 1	Seed 2	Seed 3	Seed 4
100	2	18	15	44
1000	128	102	148	110
10000	981	923	1115	899

7.6.3 Single agent, single cell observations

Coaching, in its simple hands off format, synthesised behaviours with a 10% greater chance of an action, on which the associated hypothesis is based, being selected for a given set of precepts. This is only a small bias, even when compared with the flat probability of the remaining choices, yet the results of the system behaviour are dramatically improved. However, this is a very simple example and the improved behaviour is what was intuitively expected. All that we may say of this is that in a simple system with a simple definition of influence it is possible to improve agent performance by biasing its behaviour so that it executes influential actions. We make the system slightly more complex by making it a heterogeneous system where one agent depends on another agent to enable its ability to influence the environment.

7.7 Bridge building, a two agent - single location simulation

The first simulation was an exercise in state space exploration, the coach was simply searching observations for evidence of change wholly driven by a single agent. This provided no opportunity for investigating serial influence which is one of the main motivations of this work. In the first experiment the actor agent was able to increment or decrement (as far as zero) any accumulator in its current location and, consequently, had full control over those parts of the environment that it can influence. In order to bring sequential influence to the simple single cell location we introduce a second agent type and modify the behaviour of the extant actor, which we shall call α. The new actor, β, can only initialise an accumulator, this means incrementing it from zero to 1. If an accumulator is already non zero then this agent's increment action will have no effect. α can now only increment a non zero accumulator and is, consequently, reliant on β to bring about a situation which allows it to exercise its influence giving a nested other agent construct. Adding agent type annotations to table 7.1 gives table 7.6 which characterises influence in this new setting. The abilities

Table 7.6: Influence in a two agent bridge building world

Observation	Reading	Reason	Notes
Transition from zero	Influence	Forwards to new state	α agents only
Transition to zero	No influence	Backwards from previous influence	α and β agents
Increased value	Influence	Forwards to new state	β agents only
Decreased value	No influence	Backwards from previous influence	α and β agents

that were contained in the single agent of the earlier experiment are now spread across two agents and this simple chance significantly complicates the world. From our privileged system builder's view we know that the only guaranteed way to build a bridge is for an α class agent to lay the foundation stone and for a β class agent to continue building on that foundation. Noise in the environment means that, occasionally, a β class agent will try to build where an α class agent has not laid a foundation and this will succeed. Similarly, an α class agent may try to lay a foundation on an extant bridge and environmental noise will cause this to succeed – maybe even carrying the bridge to a point where the observer sees it as good. Noise will play a major part in coaching agent observations. The single agent experiment had one action which was unambiguously influential and this influence continued with repetition.

In this version of the experiment we have two agents which must cooperate to bring about an increase in accumulator value beyond 1. The α class agent may either increment the accumulator from zero to one or decrement the accumulator. This is not a repeatedly influential action so, barring consistent noise, its results will not aggregate over time as for the increment action of the single agent example. In this case – and the requirement for this is obvious when looking ahead to the multi agent, multi location version – we need to increase the chance of the agent selecting an appropriate action *when it needs to execute that action* rather then increase the overall chance of that action. We move from *hands off* coaching to *hands on* coaching where coaching agents continually monitor behaviours and reinforce then to further strengthen observed influential behaviours.

The two agent experiments bring a different set of circumstances, enabling actions are, intuitively, the more influential of a pair. They correspond to the gateways discussed earlier and there may be a limited opportunity for executing such enabling actions. Here we a run the simulation without coaching, with one time hands off coaching and with continuous hands on coaching where the bias for a particular action is gradually increased as a coach repeatedly observes evidence of that action's influence. When coaching is used, we run each simulation a number of times preserving the coached agent's state across runs. This will indicate the effect of coaching on agents which have already acquired coach generated behaviours. This state preservation is simply accomplished by resetting the accumulator and allowing the agents to continue with their coached behaviours. We expect the agents to generate joint behaviour in the first of a chain of runs. When the accumulator has been reset then we expect the agents to quickly execute their joint, coached behaviour since they no longer need to go through the discovery phase of their operation.

7.7.1 Two agent, single cell results

Here we see that there is not a linear progression of accumulator values as in the generally similar single agent case. This is to be expected as when the accumulator becomes non zero the α class agent is biased towards decrementing. The maximum reached on a randomly chosen 10000 cycle run was 13 so the accumulator values in these uncoached, noise free runs stays close to zero.

Table 7.7: Results: two agent, single cell, noise free with no coaching

Cycles	Seed 1	Seed 2	Seed 3	Seed 4
100	2	6	0	0
1000	3	0	0	0
10000	2	0	0	0

Table 7.8: Results: two agent, single cell, noisy with no coaching

Cycles	Seed 1	Seed 2	Seed 3	Seed 4
100	1	6	16	23
1000	7	0	43	12
10000	10	24	3	4

We now change our tack slightly and begin investigating the dependency between agents rather than simple, single agent ability. Rather than different sequences of actions we are more interested in the repeatability of coached behaviours – especially for one time enabling actions. We continue with one seed value for the random number generator and repeat the sequences with agent state being maintained between accumulator resets.

Recalling section 3.2 and, in particular, the discussion of figure 3.2 (which, for convenience, is reproduced here as figure 7.4) we outlined the notion of *gateways* between agent action domains.

Intuitively the α class agent opens the gateway for β class agents by tipping an accumulator off of its zero state to a non zero that is will respond to β*'s* increment actions. The earlier *hands off* coaching certainly increased the probability of an α class agent action appropriately and this is an increased probability of an

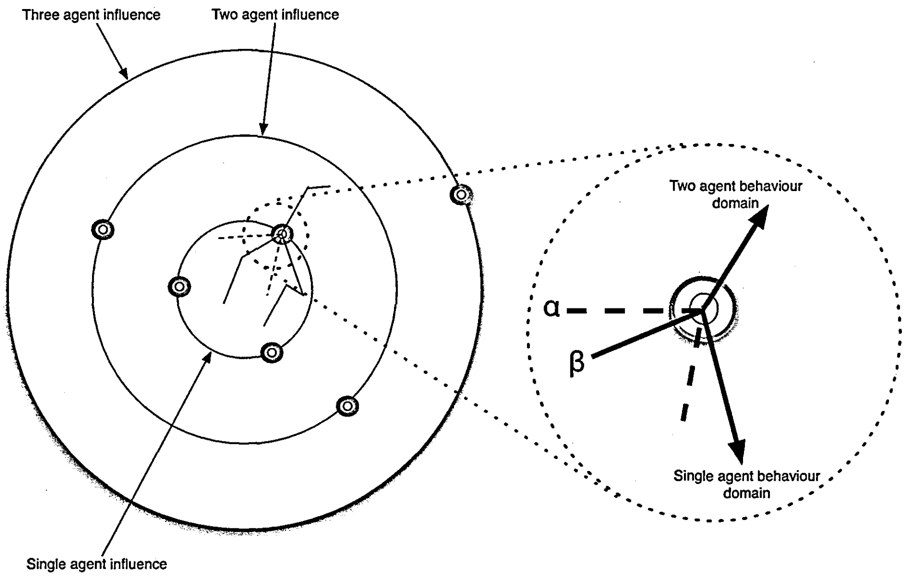


Figure 7.4: Gateways between single and multi agent influence domains (from chapter 3)

open gate but there is still a chance that the gate will not open when required. The *hands on* approach should allow a coaching agent to see which actions are influential and gradually tweak these so as to allow us to greatly increase the bias towards a particular action for a given set of percepts. Such a great increase may not be wise in a hands off coaching arrangement, noise may give the impression of an action having influence and if this verisimilar evidence leads to a large bias towards a particular action then the system's performance may be compromised.

Table 7.9: Results: two agent, single cell, noise free with coaching, 1000 cycles

Coaching	Sequence 1	Sequence 2	Sequence 3	Sequence 4	Sequence 5
Hands off	47	51	31	26	1
Hands on	340	881	866	854	887

Table 7.9 presents results from a series of 1000 cycle runs in a noise free setting using both hands on and hands off coaching. The hands off results show a significant improvement over the uncoached 1000

cycle values of table 7.7 indicating a further significant improvement over the hands off data. Introducing noise into the system gives the results displayed in table 7.10.

Table 7.10: Results: two agent, single cell, noisy with coaching, 1000 cycles

Coaching	Sequence 1	Sequence 2	Sequence 3	Sequence 4	Sequence 5
Hands off	77	75	15	67	45
Hands on	636	655	629	629	682

Focusing, now, of the *hands on* coaching approach we view coach data as raw hypotheses and trigger precepts, these are presented in table 7.11. This table has three parts, the top part indicates the patches that the coach has dropped into the environment. The centre section shows the top six patches when ordered by $P\!:C$ ratio and the bottom section shows the top six patches ordered by $P - C$ value. Hypotheses are written as $agent/choiceID \rightsquigarrow Accumulator : transitionID$. The percepts column has three numbers, the first is a boolean indicator of the presence of an accumulator and the second is a boolean accumulator is zero indicator.

7.7.2 Two agent, single cell observations

The results indicate that the hands on coaching method is much better at maximising agent influence in a noisy environment. Both of the systems complete the first two sequences with broadly similar results but the invariance of behaviour weightings based on verisimilar evidence eventually causes a degradation of system performance. The hands on coaching gradually increases the bias towards actions that are consistently observed to have influence and this has a self reinforcing effect by reducing the probability of actions which are inconsistently influential. Behaviour tweaking in this instance is positive, influential behaviours are strengthened. Non influential behaviours may, equally, be weakened and this is something which we shall address when discussing further work based on this theory of influence. This set of simple experiments takes us one step closer to a multi cell simulation where the reduced density of accumulators combined with the

Table 7.11: Patch generator data: two agent, single cell, noisy, hands-on coaching

ID	Precepts	Hypothesis	PCR	PMC
		Dropped patches		
1	10	$\beta/5 \rightsquigarrow AC1 : 3$	0.91133	3006
12	10	$\alpha/5 \rightsquigarrow AC1 : 3$	0.755633	295
3	11	$\beta/3 \rightsquigarrow AC1 : 1$	0.0299401	-157
2	11	$\alpha/5 \rightsquigarrow AC1 : 1$	0.0140562	-484
		Patch database, PCR ordering		
1	10	$\beta/5 \rightsquigarrow AC1 : 3$	0.91133	3006
12	10	$\alpha/5 \rightsquigarrow AC1 : 3$	0.755633	295
13	10	$\alpha/4 \rightsquigarrow AC1 : 3$	0.746627	329
9	10	$\alpha/6 \rightsquigarrow AC1 : 3$	0.720497	497
5	10	$\alpha/3 \rightsquigarrow AC1 : 3$	0.717201	298
11	10	$\alpha/0 \rightsquigarrow AC1 : 3$	0.713396	274
		Patch database, PMC ordering		
1	10	$\beta/5 \rightsquigarrow AC1 : 3$	0.91133	3006
9	10	$\alpha/6 \rightsquigarrow AC1 : 3$	0.720497	497
13	10	$\alpha/4 \rightsquigarrow AC1 : 3$	0.746627	329
5	10	$\alpha/3 \rightsquigarrow AC1 : 3$	0.717201	298
12	10	$\alpha/5 \rightsquigarrow AC1 : 3$	0.755633	295
11	10	$\alpha/0 \rightsquigarrow AC1 : 3$	0.713396	274

agent's ability to move (replacing most of the null actions in earlier simulations) will enormously increase the size of the state space.

7.8 Bridge building, a multi agent multiple cell simulation

The single cell, two agent world has a formidable state space but this is constrained by the agent's other choices - those not directly affecting the accumulator - not doing anything. In the multiple location environment agents are able to move from location to location and this causes the state space to expand further still providing more of a challenge for the coach to find appropriate behaviours. The single cell environment was expanded to a three by three cell torus. This is, for the most part, simply a scaled version of the single cell worlds discussed above. Cells may contain a number of agents but only one coach. If a coach attempts to move into a cell occupied by another coach then that move is blocked and the coach remains in its current cell. During each run the cells are driven at random in a round-robin manner, every cell will operate once per cycle in a randomly selected slot. The single agent experiments had agents and coaches operating sequentially with agents on a random round robin schedule. This is carried over to the multiple cell world where if there are a number of agents in a cell then they operate on a random round robin and will then be followed by any coaching agent in that cell. All agents are able to move horizontally and vertically but coach operation is slightly different, coaching agents will execute a full coaching cycle and then a move cycle rather than a single cycle where moving is one available option. This seemingly expedient move is simply to increase coaching agent's coverage of a world and is in line with our notion that coaching agents operate at a privileged level. A cell may only contain a single coach, if a coach attempts to move into a cell that is already occupied by another coach then the attempted move fails and the coach remains in its current cell (not moving is also an option for coach selected moves). Moving from a single cell to multiple cell setting increases the size of the state space enormously. In the single cell setting agents had a number of equivalent null actions, these are now proper actions which may move an agent away from a cell where it exerts influence. In a multi cell world, if there is only one cell where an agent may exert influence then successive moves may take the agent further away from that cell progressively reducing its chances of exercising its influence. In a single cell a succession of null actions do not alter the agent's chance of being able to exert influence in its next cycle. The single cell world may be classed as *episodic* and the multi cell world as a non episodic environment as per Russell and Norvig's [101] descriptions as were presented in section 2.2.2.

Our expectations are that coaching agents will, by reinforcing influential actions, guide the agents in such a manner that they are likely to maximise their influence in a given environment.

7.8.1 Multi agent multiple cell results

The environment was first explored to build an idea of how much influence raw agents have in noisy and noise free environments. This was repeated with increasing numbers of agents to gain some insight into whether or not a simple *mob handed* approach with α and β agents duplicated in numbers sufficient to make each of the cells seem like a single cell environment.

Table 7.12: Results: 3x3 world, two agents, no coaching

Cycles	Accumulator 1	Accumulator 2	Accumulator 3
		Noise free	
100	0	4	2
1000	0	0	0
10000	0	10	1
100000	0	0	0
1000000	0	0	1
		Noisy	
100	5	6	10
1000	7	4	3
10000	21	37	13
100000	147	21	10
1000000	81	4	0

The next step was to test both varieties of coaching in noisy and noise free settings.

7.8.2 Multi agent multiple cell observations

We may infer from tables 7.12, 7.13 and 7.14 that uncoached, randomly acting agents have no influence and that the results presented are similar to those of tables 7.9 and 7.10 where two agents were bounded by a single cell. Simply increasing the number of agents seems to lessen their overall influence and, in the noisy environment, dampen the effects of random accumulator noise.

Table 7.13: Results: 3x3 world, four agents, no coaching

Cycles	Accumulator 1	Accumulator 2	Accumulator 3
Noise free			
100	0	3	0
1000	1	0	0
10000	2	0	0
100000	1	0	0
1000000	0	0	0
Noisy			
100	0	0	7
1000	6	16	29
10000	3	36	17
100000	17	3	26
1000000	126	0	17

Table 7.14: Results: 3x3 world, two agents (α and β), no coaching

Cycles	Accumulator 1	Accumulator 2	Accumulator 3
Noise free			
100	3	0	4
1000	3	0	1
10000	3	1	1
100000	0	2	0
Noisy			
100	3	3	2
1000	1	2	8
10000	2	5	4
100000	0	2	2

When coaching was introduced we see weak evidence of agent influence with hands-off coaching and much stronger evidence of influence with hands on coaching in both noise free and noisy settings, tables 7.15 and 7.15 respectively.

Turning to analysis of the agent's stacks of acquired up behaviours after the noisy environment run of table 7.16 we see:

Table 7.15: Results: 3x3 world, two agents (α and β), noise free

Cycles	Accumulator 1	Accumulator 2	Accumulator 3
Hands off coaching			
100	1	5	0
1000	13	6	2
10000	1	27	4
100000	44	22	1
Hands on coaching			
100	19	1	1
1000	449	261	1
10000	2526	3085	2382
100000	3010	28775	27870

Table 7.16: Results: 3x3 world, two agents (α and β), noisy

Cycles	Accumulator 1	Accumulator 2	Accumulator 3
Hands off coaching			
100	11	4	9
1000	5	4	23
10000	83	92	33
100000	14	317	43
Hands on coaching			
100	5	0	6
1000	199	208	345
10000	2810	2137	2880
100000	28137	29141	27955

```
Agent Bingo(alpha)  Behaviour dump start.

0) ID = 1 Trigger = 11 Weighting = [57, 114, 171, 228, 285, 942, 1000]

1) ID = 8 Trigger = 10 Weighting = [7, 14, 21, 28, 985, 992, 1000]

Agent Bingo Behaviour dump end.


Agent Fleegle(beta)  Behaviour dump start.
```

```
0) ID = 0 Trigger = 10 Weighting = [7, 14, 21, 28, 35, 992, 1000]
Agent Fleegle Behaviour dump end.
```

These behaviour stacks have a number of interesting features, behaviour ID 1 captures the α class agent's ability to tip the accumulator from a zero state even in a very noisy setting. The bias weightings – with block sizes of 57 – indicate that this behaviour did not receive a full complement of reinforcing tweaks indicating that coaching agents did not observe many instances of this behaviour. This behaviour's core hypothesis has, as expected, a low tag number indicating that it was observed very early in the run. Its non zero value indicates that this was not the first influence that the coaching agent observed yet it managed to capture α's ability in the presence of heavy noise. The noisiness of the environment is evident in behaviour 8 which erroneously attributes influence to α class agents, the block size of its behaviour distribution indicates that this was a frequently observed event. Behaviour 0 correctly identifies β class agent ability, this is not such an interesting result because of the frequency of such actions.

The coaching agent maintains a behaviour patch stack separately from the hypothesis database. This is updated in tandem with the hypothesis database so that behaviours are reinforced as stronger evidence for them emerges. When the coaching agent droops new behaviours it uses the hypothesis database, with its evidence data, to select the behaviours to seed and simply copies them into the environment. The hypotheses where agents genuinely have influence have risen above the noise induced influence and α's influential behaviour – which occurs very rarely – is still seeded because it is the most influential hypothesis for a given set of trigger precepts. This is the best explanation that the coaching agent has for that influence and is the one that it seeds.

7.9 Revisiting nested other agent STIT statements

In section 6.6 we discussed operational aspects of coaching agents and, in particular, the structure of its database. Recall that the database is a series of hypothesis buckets which represent perceived states in a manner which replicates the hypothesis tree of our theory with the addition of cyclic links. Each of these state buckets contains a set of hypotheses which have links to other buckets representing achieved states

Table 7.17: Patch generator data: two agent, multiple cell, noisy, hands-on coaching

ID	Precepts	Hypothesis	PCR	PMC
		Dropped patches		
0	10	$\beta/5 \rightsquigarrow AC1 : 3$	0.884134	13122
8	10	$\alpha/4 \rightsquigarrow AC1 : 3$	0.627907	11
1	11	$\alpha/5 \rightsquigarrow AC1 : 1$	0.00050045	-9981
		Patch database, PCR ordering		
0	10	$\beta/5 \rightsquigarrow AC1 : 3$	0.884134	13122
8	10	$\alpha/4 \rightsquigarrow AC1 : 3$	0.627907	11
6	10	$\alpha/6 \rightsquigarrow AC1 : 3$	0.580645	5
9	10	$\alpha/2 \rightsquigarrow AC1 : 3$	0.555556	4
10	10	$\alpha/0 \rightsquigarrow AC1 : 3$	0.53125	2
3	10	$\alpha/3 \rightsquigarrow AC1 : 3$	0.0597125	-10780
		Patch database, PMC ordering		
0	10	$\beta/5 \rightsquigarrow AC1 : 3$	0.884134	13122
8	10	$\alpha/4 \rightsquigarrow AC1 : 3$	0.627907	11
6	10	$\alpha/6 \rightsquigarrow AC1 : 3$	0.580645	5
9	10	$\alpha/2 \rightsquigarrow AC1 : 3$	0.555556	4
10	10	$\alpha/0 \rightsquigarrow AC1 : 3$	0.53125	2
7	10	$\beta/1 \rightsquigarrow AC1 : 3$	0.00885609	-1331

and a set of evidence data. The final act of coaching agents is to log their database structure before closing down. Casting the data of table 7.17 into this format gives us the structure of figure 7.5. Does this capture a nested other agent sequence? Recall that in section 7.7 we stated that the two agents have different abilities, α agents can only increment an accumulator from zero to one We see that from database bucket 0 which holds the state equivalence chain for 11 percepts which indicate that the accumulator is zero, the strongest path through the database structure to state bucket 1, where the accumulator is non-zero, which admits α and β agents is $\alpha/5$ followed by $\beta/5$, both indicated by heavier lines in figure 7.5. The first $\alpha/5$ takes the system from state 11 to state 10 and subsequent $\beta/5$ actions increase influence. Although the environment was noisy the only transition with observed agent involvement that takes the system from a zero accumulator state to a non zero accumulator state is $\alpha/5$. All of the β agent driven transitions – ignoring the question of noise from

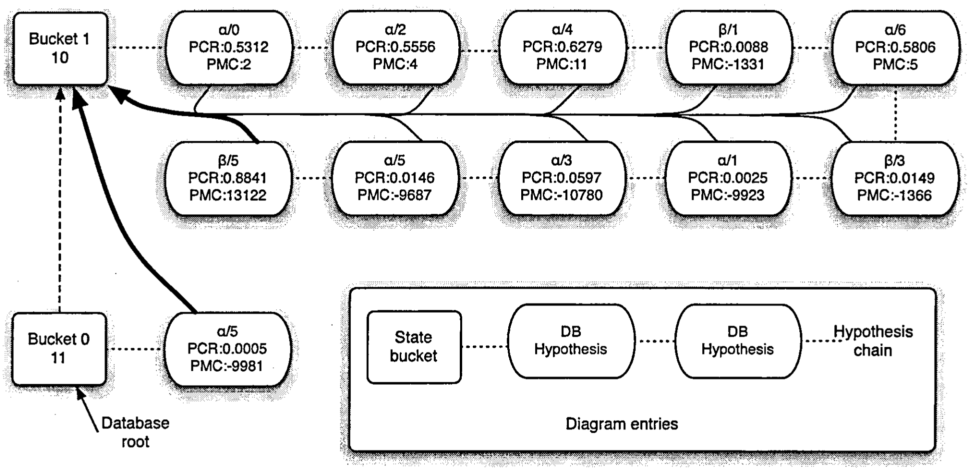


Figure 7.5: Database structure generated by a simple experiment

the moment – take the system from a non-zero accumulator state to a non-zero accumulator state. Granted, this is a very simple system and the coaching agent's concept of influence may be construed as favouring this result. It is, however, a very noisy environment and the coaching agent's notion of influence was deliberately made simple so as to allow for manual data inspection. Even with a simplistic view of the world the only agent driven path from a state where the accumulator is zero is identified. Combining this with the loop at bucket 1 we see that there is a path through the database records where α can influence a transition to bucket 1 where β may repeatedly exercise its influence. This structure captures a situation where nested other agent influence is displayed and we feel that it is fair to say that [α *influences*: [β *influences*: A]] and that an influence operator provides a viable reading for STIT which admits other agent nesting with no system overhead and without compromising individual agency. Influence nested in such a way is not absolute in the sense that a strict STIT is but this absolute nature is present if both agents choose to act appropriately.

In earlier investigations leading to this work we considered the notion of agents influencing other agents in such a way as to lead to complex behaviours. We envisaged a tiered structure, figure 7.6, with simple, single agent behaviours at the bottom and with increasingly complex behaviours on top. Our question then was what was required to jump the gaps between behaviour levels and initially we dubbed this as some form of magic which we hoped to observe and characterise in such a way as to alloy synthesis of new behaviours. The initial hypothesis was that this magic was data mining and this is an avenue that was explored. In later

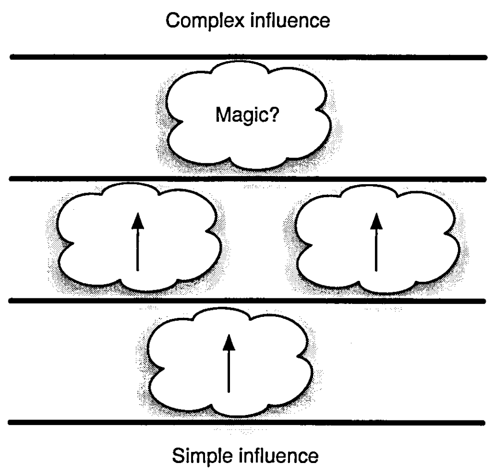


Figure 7.6: Complexity and magic

stages of the research it became evident that the progression was not really a progression of levels as in figure 7.6 and was more accurately depicted as a partitioning of a world state space, as in figure 7.4 with the magic being the opening of gateways to other areas. This is captured in the hypothesis and database structure. Our future investigations will confirm if this is adequate for richer environments or if we need to have data mining of the hypothesis structure to synthesise more complex behaviours.

7.10 Reviewing experimental results

The experiments described above illustrated a progression of complexity from trivial to complex. Whilst the settings may have been simple the introduction of noise made the job of coaching agents significantly more difficult, particularly in later experiments where one agent class carried out a gateway action by potentially masking gateway actions. Similarly, the choice of what constitutes influence may appear as if it is intended to deliberately steer the agents towards a goal. This is not the case, the choice of what constitutes influence is intended to guide agents towards exploring their state space, new values are influential and previously seen values are not. The choice of what does and what does not constitute influence is accumulator value agnostic.

The single agent, single cell experiments behaved as expected in noise free settings. These single cell noise free experiments are, perhaps, thought of as test runs to verify that the experimental software operated as expected. Because of the single cell nature of the world actions which would have caused an agent to move away from an accumulator are treated as null actions. The same experiments with noise introduced produced rather better than expected results with coached agents performing almost as well as they did in a noise free setting.

The two agent single cell location introduced a complication in that it required sequencing of actions, this is the notional gateway action introduced in chapter 3. The two agent single cell simulation also saw the introduction of the notions of "hands off" and "hands on" coaching. This indicated that the hands on approach, where the coaching agent continually monitors actor agent influence, allows the system to overcome the effects of noise which may manifest itself as apparent influence.

Moving to multi cell simulations we see that world becoming more complex with a significant increase in state space size and the introduction of dependencies, α class agents can only increment accumulators with a zero value and β class agents can only increment non zero accumulators. Agents can now move from cell to cell, this replaces some of the agent's null actions leaving only one null in its choice set. This reduced the impact of noise on non influential actions slightly because null actions in the presence of an accumulator may be replaced by actions moving an agent away from an accumulator. This slight "weakening" of the effects of noise was counteracted by the increased complexity of the environment. The density of accumulators was greatly reduced with a concomitant reduction of the probability of coaching agents observing the gateway action where an α class agent increments a zeroed accumulator. Without coaching agents were, essentially, ineffective and hands off coaching did not provide any great improvement. Hands on coaching, in this setting, provided impressive results with the system identifying the gateway action, illustrated in figure 7.5, and developing agent behaviour patterns that seemed well suited to achieving the observer's goals.

This is a strong result which potentially gives a starting point for further research into the allocation of behaviours amongst agents. Intuitively, if a β class agent had two increment actions, one for zero accumulators and one for non zero accumulators, then the system performance may have been poorer. One possible approach would be a two stage coaching process, a hands off stage allowing for the identification of influential actions, an agent classification where certain agent actions are suppressed leaving the influential actions (with some others) followed by a hands on coaching stage. One of the early questions in this research was the difference between what ought to be done and what agents ought to do and this seems to point to that question and offers a niche – albeit with simple agents in a simple environment – which may provide insights in future exploration.

Chapter 8

Discussion and further work

There are a number of threads that run through this work, the deontic identity that *ought implies can*, Milner's statement that the behaviour of a system is just what is observed and Chellas's observation that it would be bizarre to deny the sense of nested STIT. All three of these are drawn together by the notion of influence, ought implies can and Milner's statement may be grouped by our working with observations. If a system does not have the raw, stochastic ability to bring about A then we will not observe it bringing about A and it would be senseless for an observer to expect it to bring about A. Milner's statement and Chellas's observation may be grouped because by observation we may see patterns of behaviour where one agent's choices have some effect on another agent's behaviour. Chellas's observation and the deontic identity may be grouped because when a coaching agent observes instances of one agent influencing another it attempts to maximise that influence.

Ought implies can requires that a proposition or state be *accessible* and this is where we laid the foundations of our notion of agent influence. We outlined the notion of accessibility in section 3.2 where we introduced the idea of viewing a system's state space as a set of influence domains. These domains extend from single agent influence to an arbitrary degree of complexity where only multiple agents may have influence, this was illustrated in figure 3.2. Our investigation led us to note some of the difficulties that standard STIT semantics present when dealing with other agent constructs. There are alternative readings for STIT which address these difficulties but these these also bring significant overheads which compromise their usability with very simple agents. We have introduced the notion of influence as an alternative to these readings and have indicated how our view of influence differs from others, notably the work of Ferber et al. and their notion of centrally aggregated influence. We have developed a theory of influence and have offered a partial logical characterisation of our influence operators to indicate that they are similar to STIT operators. We have carried out simple experiments to indicate that the notion of influence is viable in practical applications. This leaves us in the position where we are able to discuss what has been achieved so far, indicate where our theory of influence may be useful and outline where we intend to work so as carry our influence theories further.

This leads us to a point where we can discuss approaches to coaching. So far we have simply considered the process of detecting agent influence and have not really said much about what to do with it – other than attempt to increase it – when we have detected it. The series of experiments above provide some clear examples of what we term "hands off" and "hands on" coaching. The hands off approach is simply where a coach detects influence, seeds a behaviour which increases the probability of this influence occurring for a given set of precepts – in this case it was an arbitrary 10% increase – and leaves it at that. The hands on approach is where a coach detects influence, seeds a behaviour as in the hands off approach but continues to monitor that behaviour and increases the bias when it detects future instances of its having influence. Consider the gateways between influence domains, if two agents meet then ideally they should immediately act in such a way that allows one or both of them to create or pass through a gateway to the next influence domain. An arbitrarily small increase in the chance of each agent selecting the correct action from a potentially large set of actions is clearly not good, something approaching certainty would be much better here. Certainty of action is something which we deliberately steer away from, our entire approach to synthesising new behaviours is founded on our ability to observe them and if agent behaviour becomes fixed then the chance of randomly generated new behaviours disappears. This is where our view of the ought implies can identity differs from the more standard point of view.

Ought implies can represents stochastic ability and this simply says that a system is somehow capable of, for example, bringing about A. Our coaching agent's analysis of influence is an attempt to turn the "can" from a stochastic ability to a set of behaviours that not only guarantees A, but which can also be associated with an agent or set of agents. Returning to our leads to and may lead to operators, the "can" in our system is represented by links between coach database buckets and the "ought" is the combination of detected influence and agent behaviour weighting. We conclude by making some observations based on this work, outlining what we see as its limitations, outlining where this work will be carried in the future and by making some tentative suggestions for applications of our notion of influence.

8.1 Observations

The initial exploration and set of simple experiments, described in chapter 3, provided encouraging results. Agents are certainly able to "carry" an undischarged influence through a noisy environment. They may do this by, for example, holding a token or tool of some sort and not using it until appropriate. We were able to "track" such instances and by aggregating observed histories build a representation of the world which identified where an undischarged influence was "created". We followed this exploration by transferring our attention to a practical setting and deliberately introduced noise into our experiments. The final experiment, although set in a simple environment, provided encouraging results. Despite noise and the very small incidence of $\alpha/5$ we were clearly able to identify a chain of events leading to an influence extending, bridging like behaviour. The detection of influence and the representation of serial influence, with one agent's ability being contingent on another was due, in part, to the simplicity of the system. However, the noise content of the experiments makes the results of chapter 7 look impressive given the masking effect that the injected noise will have had on agent influence.

Our partial characterisation indicates that, in common with STIT, influence supports modal operators. More importantly the characterisation indicates that our theory of influence may be extended into domains requiring complex sequences of agent actions. The failed characterisations are equally important, the other agent extension indicated that our notion of influence rejects cases where agent actions may be mutually exclusive. These results bode well for further investigation to build a more solid understanding of exactly how a coach may manipulate evidence of uncertain ability without generating unrealistic conclusions.

8.2 Summary of contributions

The contributions of this work were listed briefly in section 1.2 and we revisit this list so as to illustrate how the contributions fitted together to build a coherent system.

The notion of influence allowed the development of a theory intended to overcome some of the shortcomings of STIT semantics. This was approached with a particular view to managing multi agent ability where overall ability is dependent on aggregate or sequential agent actions.

The notion of *extended* influence provided a foundation allowing the notion of single agent influence to be extended into scenarios involving a number of agents. This extension into the domain of multiple agent actions does not compromise individual agency and allows the notions to be applied to simple, reactive agents without requiring any extra societal overhead.

The notion of gateway actions allowed the characterisation of this approach as a state space search. This provided insights into agent behaviour and, in particular, the *ought implies can* deontic identity.

The notion of strict STIT() provided a means of differentiating this influence based approach from standard STIT() allowing the development of new operators which would allow for a logical characterisation.

The modal *leads to* and *may lead to* operators which pave the way for the treatment of the notion of influence in a similar manner to standard stit expressions.

The partial logical characterisation of *leads to* and *may lead to* operators supported their use in inferences of agent influence allowing the development of a practical system to illustrate the use of influence in a nested behaviour setting.

The binary representation of agent choice allowed the representation of hypotheses of agent influence in a simple and computationally tractable manner which allowed the representation of the evolution of an agent world.

The application of discrete time to the standard branching time framework allowed the computational problems associated with the unbounded nature of branching time to be addresses.

8.3 Limitations of this work

We have taken a simplistic view of agent influence in this work. We have tended to focus on an agent's ability to change its environment. We have also examined only simple agent and action combinations, coaching agent hypotheses are predicated on the selection of a single action from an agent's set of actions. This was, as we mentioned in section 7.9, a deliberate choice to allow for manual data inspection. Our next step is to gradually introduce complexity into the system, perhaps having accumulators change state after a certain point and then respond to a different set of actions. Gradual enrichment of the environment will lead us to an understanding of subtleties of agent interaction which will place us in a position to introduce automated data mining and apply the influence concept to complex systems.

8.4 Future work

This work has examined influence in a very narrow setting, that of agents with well defined choice partitions and a limited number of available actions. This has allowed us to observe that influence is a measurable and manageable feature of agent systems. An obvious next step is to carry the theory and experimental work into more complex domains, domains where complexity becomes the major obscuring factor rather than noise. In order to do this we will need to extent our research into supporting techniques and research areas and some of these are listed below.

8.4.1 Further characterisation

STIT has been the subject of extensive research and has generated a large amount of literature. We offer influence as a potential alternative allowing the extension of STIT semantics into the other agent domain but do so guardedly. Our work on characterising influence needs to be extended so that we may identify where STIT and influence truly differ and focus investigation on these areas. We have sufficient characterisation to allow simple systems, this will need to be extended so that we may investigate systems where aggregate behaviours are nested within each other leading to very complex behaviours.

8.4.2 Heuristics for behaviour selection

The relative gap heuristics which we adopted, outlined in section 6.6.5, were based on the ratio and absolute difference between positive evidence and counter evidence tallies for a given hypothesis. These heuristics reflected observations and are viable in the setting of the simulations presented in this work. The appropriateness here comes from the fact that the experiments were carried out on a completely synthetic system and that as the builders of this system we understand the salient points of its behaviour. In more complex systems or systems that are not the synthetic product to a single entity then the salient points of system behaviour may not be known. In such cases the relative gap heuristics may be rather too blunt an instrument which will be unable to identify subtler aspects of system behaviour. Relative ordering of hypotheses may, for example, be taken into account along with gaps.

In complex worlds it may be that from a given state there are a number of influential actions. We noted, in section 7.5, that actor agents will hold one set of choice partition weightings for each precept set. In complex environments we expect that agents belonging to the same base class will eventually turn into different classes of coached agents, each using its abilities to different ends. At present we have an intuitive idea of how this will work but need to investigate the notion further. Of particular interest is whether or not coaching agents should specialise in coaching particular classes of coached agents or if individual coaching agents should hold what are, essentially, competing hypotheses founded on the same set of precepts.

Our approach deals with agents in the abstract by treating them as choice partitioning mechanisms. The *size* of a choice partition may provide a heuristic for evaluating other agent hypotheses on the intuition that if, for example, β/L influences α's ability at some other choice then the size of β's choice partition may be used in the coach database structure to gauge the strength of evidence for joint action. The gap heuristics may be dynamic, over longer runs a coach will see what actions have influence in certain sets of circumstances and may be able to treat the agent as a dynamically sized choice partitioning mechanism.

8.4.3 Potential use of data mining

We have alluded, at the end of the previous chapter, to the possible advantages of data mining in this research and have investigated applications of data mining in this setting, Logie et al. [83]. At this stage of our research we were happy to see that mining techniques were viable we are not data mining specialists and because of this we relied on simpler tools in the experiments of chapter 7. In simple environments where agents are guided by a small number of norms then simply examining the coaching database structure may be sufficient for identifying joint behaviours. If the system has multiple norms or involves multiply nested behaviours then tracing becomes more complex. Longer time periods between bounding events may allow a greater number of possible conclusions making identification a more demanding proposition. We believe that this may form a pattern that is fractal in nature and this is something that we intend to investigate in more complex systems with the assistance of data mining techniques.

Larger systems will have a number of coaches and this suggests some form of distributed data mining (DDM). Distributed data mining typically, Kargupta et al. [70] note, involves generating a global model and this runs counter to our multi agent approach which precludes centralisation. DDM generally allows for either multiple mining entities on a single repository or multiple entities mining remote, independent repositories. Our system does not really fit into either of these categories. It certainly has a single data repository, the data are distributed throughout the environment, but the data carries implicit location data and the mining agents are unable to see the entire data repository. Moreover the implicit location data, its dynamic nature and our agent model, following Kulkarni et al. [76], forces the miners to process data "on the fly" but leaving the environment unchanged so that other mining agents potentially following different paths will see consistent data trails.

An approach being considered is dynamic stream mining, Coaching agents are effectively following a stream of patch data and as the system approaches optimal behaviour the contents of agent data patches will change to reflect improvements, this is the dynamic nature of the data suggested above. Cohen [32] describe a system for mining a continuous dynamic data stream using an incremental single pass algorithm. The bounding events for "good" behaviours may be treated as periodic elements in the data stream that a coach is following and the rate of changes in the events between these bounds may give an indication that the agents in the system are approaching optimal behaviour.

Klusch [74] notes that that cooperation among distributed data mining processes may allow effective mining even without centralised control making this approach attractive in a multi agent setting.

8.4.4 Temporal considerations

Our treatment of the time aspect of agent action has been simplistic. Time clearly plays an important part in agent interactions and this is especially so in time critical sequences. Recent work by Tulenheimo [115] on a logic of time division $\mathcal{L}_{TD}$ indicates the need for investigation into suitable interpretation of time division for structures like trees. Our simplistic approach, with an implicit handling of time due to our characterisation of actions as being instantaneous, is adequate for the environments that we have described. If time plays a more prominent role, perhaps due to the results of an agent's action being delayed, then we will need some mechanism for factoring this into hypotheses. An approach that works with tree structures may be viable and may possibly be implemented by adding properties to links in the coaching agent's database structure.

8.4.5 More complex environments

Agents in our system are very simple and this simplicity means that they have no explicit communications system. This may be adequate for agents located in the same physical or virtual location but this precludes the investigation of influence in environments where agents are located in separate but connected environments - mobile web agents for example. Communications between agents in such settings is necessary. The difficulty here is having some communications system or language that allows coaching agents to interpret communications in a way that is similar to its view of the world. Our work is centred on social rather than individual agency and this is something that Singh [108] notes is lacking in some approaches to agent communication languages. Erdur and Seylan [39] describe an agent communication language where belief and intention may be managed. This may be extended so as to include potential ability so that agents may communicate information which allows the synthesis of joint abilities.

8.5 Possible applications

The experiments described were simple but despite this simplicity they posed problems because of the deliberately introduced noisiness of the experimental environments. This allowed us to investigate the notion of influence in something approaching a "real life" setting with noise and uncertainty. These will, typically, be areas where an agent of some sort has an ability but that the "dimensions" of that ability are unknown. Our experimental coaching agents know *a priori* of agent choice sets but this need not necessarily be the case. An image of agent ability may be built as observations are gathered and – by comparison of achievable states – the true nature of an agent's ability. We tentatively offer suggestions for potential application of our notion of influence. Some of these areas, which have arisen from discussion during this research, are listed below.

8.5.1 Pharmaceutical trials

Drug trials are complex undertakings, the environment is noisy, there are many influential factors in addition to the product under investigation. Yardimci [126] notes that so-called "soft computing" technologies have been applied widely in medicine. Our notion of influence may be thought of as a soft computing approach. Our approach is based on observation and does not have the same agent behaviour systems – mutation and crossbreeding – of genetic and evolutionary approaches. Our experiments were, despite their simplicity, challenging because of the noisiness of the experimental environment. This makes our technique useful for identifying influence in noisy and uncertain environments and it may be a useful adjunct to statistical analysis of pharmaceutical trials where the goal is to identify influential treatments.

8.5.2 Personnel management

In this work we dealt with a hard partitioning of choices. Agents were assumed to have a "hard set" of behaviours and hypotheses were based on an agent/choice pair. This need not necessarily be the case, an agent's choices may be a fuzzy set and the agent's choice mechanism need not be fully understood. Personnel management implies that the agents are human, notoriously complex cognitive agents. Such agents – despite their cognitive abilities – may be no more aware of their choice sets than our simple reactive agents. In many cases human agents "just do what seems right" without explicit consideration. The application of soft computing to business and management problems seems to be a relatively young area of research which has a small but growing body of literature, the approach of observing the influence of "agents" appears to be a novel application of computing in this area.

8.5.3 Investigation of emergent behaviour

Emergent behaviour is one of the topics discussed in this work. We stated, in section 2.12, that we make no claims for new insights into the this phenomenon. Our approach does, however, guide state space searches and may be viewed as an accelerator for emergent behaviour in simulation experiments. In conjunction with a developing logical characterisation of influence we may be able to gain insights into the mechanisms of systems exhibiting such features. Li et al. [81] note that in agent systems emergent behaviour may be deduced from the sets of actions that agents may perform but that this becomes difficult and unpredictable where agents have many choices. Our approach may constrain searches for such behaviours making potentially predictive analysis of complex systems potentially feasible. Li et al. also mention that the logic of systems should be simple in order for it to exhibit emergent behaviour. If this were the case then hybrid systems using emergent behaviour would not be possible, we feel that continued work on the characterisation of influence may provide insights into emergent behaviour.

Agent systems are attractive for exploration in hazardous environments, NASA has a prospective mission to explore asteroid belts using small autonomous craft. Rouff et al. [98] describe research on what is called an ANTS (Autonomous Nano-Technology Swarm) mission and note that this has led to a need for formal methods for swarm based systems. One of the problems with swarm systems is that testing may be extremely difficult because the system's multi agent nature means that testing is, essentially, a state space search and the state space may grow exponentially with the number of agents. Identifying gateway actions, as outlined in section 3.2, may provide a tool for constraining the state space allowing larger systems to be tested and verified.

Appendix 1 – proofs and lemmas

Proofs and lemmas for chapter 3

Backwards monotony

Definition 48 *Backwards monotony, from Belnap et al. [8, page 272]:* $w_1 \leq w_2$ *and* $m_1 \equiv^{\alpha}_{w_2} m_2$ *implies* $m_1 \equiv^{\alpha}_{w_1} m_2$

The witness identity lemma

Definition 49 *The witness identity lemma, from Belnap et al. [8, page 272] after Chellas [31]: Suppose that* Q_1 *implies* Q_2*, that m,* w_1 *and* w_2 *are moments and that* α_1 *and* α_2 *are agents, possibly identical or possibly distinct. Suppose, further, that* w_1 *is a witness to* $[\alpha_1\ stit\colon Q_1]$ *at m and that* w_2 *is a witness to to* $[\alpha_2\ stit\colon Q_2]$ *at m. Then* $w_2 \leq w_1$.

The second witness identity lemma

Definition 50 *The second witness identity lemma, from Belnap et al. [8, page 272]: If w is the witness to to* $[\alpha\ stit\colon Q]$ *at* m_1 *and if* $m_1 \equiv^{\alpha}_{w} m_2$ *then w is also the witness to* $[\alpha\ stit\colon Q]$ *at* m_2 *– which is therefore settled true at* m_2.

Proof of the impossibility of $[\alpha\ stit\colon [\beta\ stit\colon A]]$

From [8, page 274].

(a) Assume that $[\alpha\ stit\colon [\beta\ stit\colon A]]$ is settled true at m_1 with w as a witness and m_2 as a counter satisfying the negative condition so that:

(b) $[\beta\ stit\colon A]$ is not settled true at m_2.

By independence of agents there must be an m_3 such that both:

(c) $m_1 \equiv_w^\alpha m_3$

and:

(d) $m_3 \equiv_w^\beta m_2$

By (a), (c), and the second witness identity lemma it must be that:

(e) w is a witness to $[\alpha\ stit\colon [\beta\ stit\colon A]]$ at m_3. By (a) and (c) we must, by the positive condition, have:

(f) $[\beta\ stit\colon A]$ settled true at m_3 – let w_1 be the witness for this. From (e), (f), and the witness identity lemma, we infer:

(g) $w_1 \leq w$.

So (d) and (g) imply by backward monotony that:

(h) $m_3 \equiv_{w_1}^\beta m_2$. But then the second witness identity lemma with (f) and (f) gives that $[\beta\ stit\colon A]$ must be settled true at m_2 which contradicts (b) and completes the proof.

Appendix 2 – History traces for chapter 3

Each state, represented as in figure 8.1, contains a complete representation of the simple agent world. Each state has an ID number below the bottom, left cell. ID numbers are sequenced so that larger numbers indicate later states. Each cell contains data on its contents, α and β indicate that these agents are present in the cell.

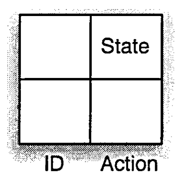


Figure 8.1: History trace key

β*'s* ability is dependent on α*'s* having given it a one use token and as a convenience this is indicated by an asterisk. Where β is "enabled" it is represented as β*. Where A holds in a cell this is written as A, if A does not hold then it will not be written in that cell. Note that A may appear and clear spontaneously.

Agent actions are indicated below the bottom right cell. These actions are the actions that *led* to that state. The available actions re N, S, E and W for move directions. Actions are always listed α first then β although they occur simultaneously. G, for α only, indicates α*'s* attempt to give a token and U, for β only indicates β*'s* attempt to use a token. Note that β may still select action U when not in possession of a token. For example, in figure 8.2 state 17, the GE indicates that at state 12 α attempted a give action and β moved east, these actions led to state 17.

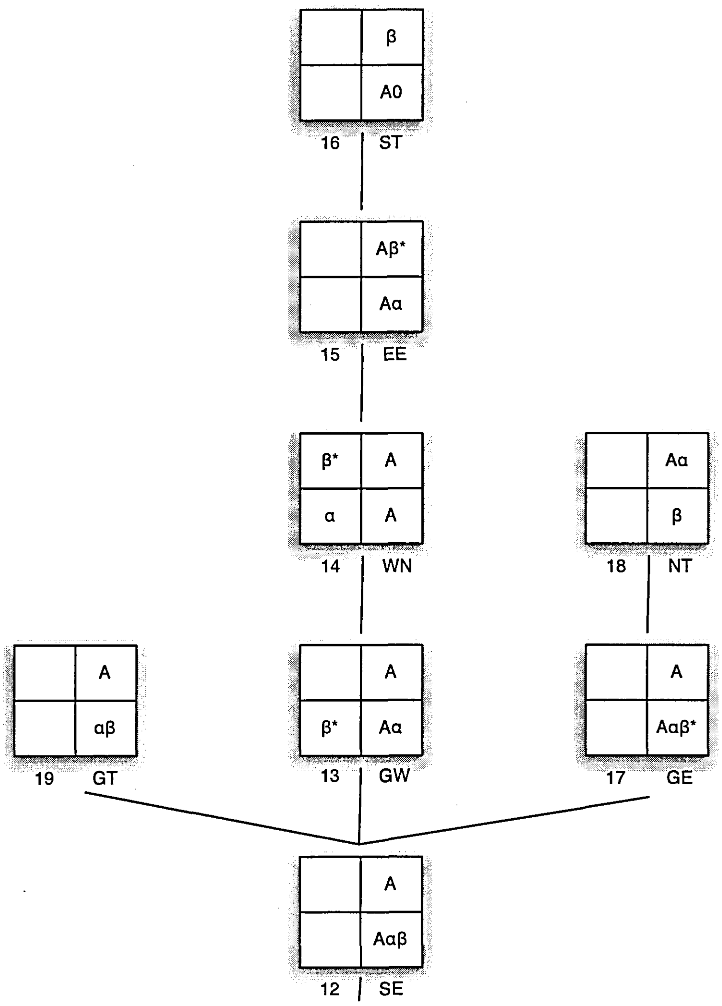


Figure 8.2: Two agent history 1: $\neg A$ brought about in three ways

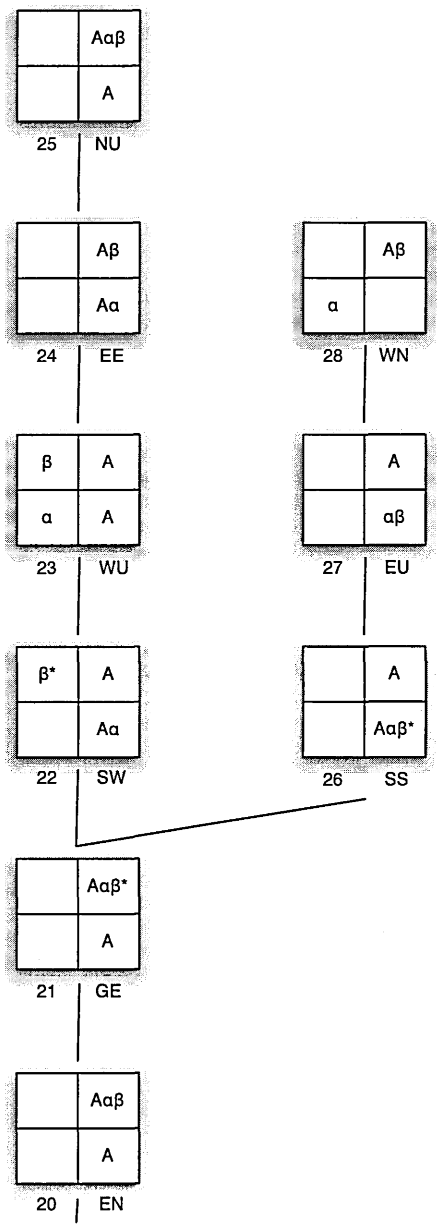


Figure 8.3: Two agent history 2: $\neg A$ brought about and $\neg A$ not brought about

Appendix 3 – tables

Table 8.1: Potential serial influence with coached α, global hypotheses ranked by $P-C$ value

Hypothesis	Evidence			P:C ratio	$P-C$ value	Sound
	Positive-	Negative-	Counter-			
$GH\beta/L \leadsto Y$	253	✓	11	0.958333	242	✓
$GH\beta/L \leadsto X$	366	✓	162	0.693182	204	?
$GH\gamma/M \leadsto Z$	212	✓	14	0.938053	198	✓
$GH\alpha/K \leadsto Z$	191	✓	50	0.792531	141	✗
$GH\gamma/M \leadsto Y$	182	✓	44	0.80531	138	✗
$GH\alpha/K \leadsto Y$	188	✓	53	0.780083	135	✗
$GH\alpha/K \leadsto X$	178	✓	63	0.738589	115	?
$GH\gamma/M \leadsto X$	151	✓	75	0.668142	76	✗
$GH\beta/L \leadsto Z$	0	✗	0	0	0	✗

Table 8.2: Potential serial influence with coached α, X prefixed hypotheses ranked by $P-C$ value

Hypothesis	Evidence			P:C ratio	$P-C$ value	Sound
	Positive-	Negative-	Counter-			
$X1\beta/L \leadsto Y$	157	✓	6	0.96319	151	✓
$X1\gamma/M \leadsto Z$	136	✓	9	0.937931	127	✓
$X1\gamma/M \leadsto X$	131	✓	14	0.903448	117	×
$X1\beta/L \leadsto X$	139	✓	24	0.852761	115	?
$X1\gamma/M \leadsto Y$	123	✓	22	0.848276	101	×
$X1\alpha/K \leadsto Y$	126	✓	25	0.834437	101	×
$X0\beta/L \leadsto Y$	96	✓	5	0.950495	91	✓
$X1\alpha/K \leadsto X$	121	✓	30	0.801324	91	?
$X1\alpha/K \leadsto Z$	118	✓	33	0.781457	85	×
$X1\beta/L \leadsto Z$	123	✓	40	0.754601	83	×
$X0\gamma/M \leadsto Z$	76	✓	5	0.938272	71	✓
$X0\beta/L \leadsto Z$	80	✓	21	0.792079	59	×
$X0\alpha/K \leadsto Z$	73	✓	17	0.811111	56	×
$X0\gamma/M \leadsto Y$	59	✓	22	0.728395	37	×
$X0\alpha/K \leadsto Y$	62	✓	28	0.688889	34	×
$X0\alpha/K \leadsto X$	57	✓	33	0.633333	24	?
$X0\gamma/M \leadsto X$	20	✓	61	0.246914	-41	×
$X0\beta/L \leadsto X$	24	✓	77	0.237624	-53	?

Table 8.3: Potential serial influence with coached α, Y prefixed hypotheses ranked by $P - C$ value

Hypothesis	Evidence			P:C ratio	$P - C$ value	Sound
	Positive-	Negative-	Counter-			
$Y1\beta/L \rightsquigarrow Y$	205	✓	7	0.966981	198	✓
$Y1\gamma/M \rightsquigarrow Y$	164	✓	13	0.926554	151	✗
$Y1\gamma/M \rightsquigarrow Z$	164	✓	13	0.926554	151	✓
$Y1\alpha/K \rightsquigarrow Y$	167	✓	18	0.902703	149	✗
$Y1\alpha/K \rightsquigarrow X$	165	✓	20	0.891892	145	?
$Y1\beta/L \rightsquigarrow Z$	163	✓	49	0.768868	114	✗
$Y1\alpha/K \rightsquigarrow Z$	145	✓	40	0.783784	105	✗
$Y1\gamma/M \rightsquigarrow X$	136	✓	41	0.768362	95	✗
$Y1\beta/L \rightsquigarrow X$	142	✓	70	0.669811	72	?
$Y0\gamma/M \rightsquigarrow Z$	48	✓	1	0.979592	47	✓
$Y0\beta/L \rightsquigarrow Y$	48	✓	4	0.923077	44	✓
$Y0\alpha/K \rightsquigarrow Z$	46	✓	10	0.821429	36	✗
$Y0\beta/L \rightsquigarrow Z$	40	✓	12	0.769231	28	✗
$Y0\beta/L \rightsquigarrow X$	21	✓	31	0.403846	-10	?
$Y0\gamma/M \rightsquigarrow Y$	18	✓	31	0.367347	-13	✗
$Y0\alpha/K \rightsquigarrow Y$	21	✓	35	0.375	-14	✗
$Y0\gamma/M \rightsquigarrow X$	15	✓	34	0.306122	-19	✗
$Y0\alpha/K \rightsquigarrow X$	13	✓	43	0.232143	-30	?

Table 8.4: Potential serial influence with coached α, Z prefixed hypotheses ranked by $P - C$ value

Hypothesis	Evidence			P:C ratio	$P - C$ value	Sound
	Positive-	Negative-	Counter-			
$Z1\beta/L \rightsquigarrow Y$	198	✓	8	0.961165	190	✓
$Z1\beta/L \rightsquigarrow Z$	186	✓	20	0.902913	166	×
$Z1\alpha/K \rightsquigarrow Z$	177	✓	13	0.931579	164	×
$Z1\gamma/M \rightsquigarrow Z$	167	✓	10	0.943503	157	✓
$Z1\gamma/M \rightsquigarrow Y$	145	✓	32	0.819209	113	×
$Z1\alpha/K \rightsquigarrow Y$	147	✓	43	0.773684	104	×
$Z1\alpha/K \rightsquigarrow X$	143	✓	47	0.752632	96	?
$Z1\gamma/M \rightsquigarrow X$	121	✓	56	0.683616	65	×
$Z0\beta/L \rightsquigarrow Y$	55	✓	3	0.948276	52	✓
$Z0\gamma/M \rightsquigarrow Z$	45	✓	4	0.918367	41	✓
$Z1\beta/L \rightsquigarrow X$	123	✓	83	0.597087	40	?
$Z0\alpha/K \rightsquigarrow Y$	41	✓	10	0.803922	31	×
$Z0\gamma/M \rightsquigarrow Y$	37	✓	12	0.755102	25	×
$Z0\beta/L \rightsquigarrow X$	40	✓	18	0.689655	22	?
$Z0\alpha/K \rightsquigarrow X$	35	✓	16	0.686275	19	?
$Z0\gamma/M \rightsquigarrow X$	30	✓	19	0.612245	11	×
$Z0\beta/L \rightsquigarrow Z$	17	✓	41	0.293103	-24	×
$Z0\alpha/K \rightsquigarrow Z$	14	✓	37	0.27451	-23	×

Bibliography

[1] *The Second International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS 2003, July 14-18, 2003, Melbourne, Victoria, Australia, Proceedings*. ACM, 2003.

[2] T. Ågotnes, W. van der Hoek, and M. Wooldridge. Robust normative systems. In *AAMAS '08: Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems*, pages 747–754, Richland, SC, 2008. International Foundation for Autonomous Agents and Multiagent Systems.

[3] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2):510–530, 1985.

[4] A. R. Anderson. A reduction of deontic logic to alethic modal logic. *Mind*, 67:100–103, 1958.

[5] K. J. Åström. Optimal control of Markov decision processes with incomplete state estimation. *Journal of Mathematical Analysis and Applications*, 10:174–205, 1965.

[6] N. Belnap. Backwards and forwards in the modal logic of agency. *Philosophy and Phenomenological research*, 51:777–801, 1991.

[7] N. Belnap. Branching histories approach to indeterminism and free will. http://philsci-archive.pitt.edu/archive/00000890/00/bh-approach-indet-free-will.pdf, November 2002.

[8] N. Belnap, M. Perloff, and M. Xu. *Facing the Future: Agents and Choices in Our Indeterminist World.* Oxford University Press, Oxford, 2001.

[9] B. Bennett, C. Dixon, M. Fisher, E. Franconi, I. Horrocks, U. Hustadt, and M. de Rijke. Combinations of modal logics. *Artificial Intelligence Review*, 17(1):1–20, 2002.

[10] P. Blackburn. Modal logic as a dialogical logic. *Synthese*, 127:57–93, 2001.

[11] G. Boella and R. Damiano. An architecture for normative reactive agents. In K. Kuwabara and J. Lee, editors, *Intelligent Agents and Multi-Agent Systems, PRIMA 2002*, pages 1–17. Springer-Verlag, 2002.

[12] G. Boella, J. Hulstijn, and L. W. N. van der Torre. Virtual organizations as normative multiagent systems. In *HICSS*. IEEE Computer Society, 2005.

[13] G. Boella and L. W. N. van der Torre. Attributing mental attitudes to normative systems. In *AAMAS* [1], pages 942–943.

[14] G. Boella and L. W. N. van der Torre. Obligations as social constructs. In A. Cappelli and F. Turini, editors, *AI*IA*, volume 2829 of *Lecture Notes in Computer Science*, pages 27–38. Springer, 2003.

[15] G. Boella and L. W. N. van der Torre. Fulfilling or violating obligations in normative multiagent systems. In *IAT*, pages 483–486. IEEE Computer Society, 2004.

[16] M. Bowling. Convergence problems of general-sum multiagent reinforcement learning. In *Proc. 17th International Conf. on Machine Learning*, pages 89–94. Morgan Kaufmann, San Francisco, CA, 2000.

[17] R. Bradley and N. Swartz. *Possible Worlds: Introduction to Logic and Its Philosophy*. Blackwell, 1979.

[18] M. E. Bratman, D. Israel, and M. Pollack. Plans and resource-bounded practical reasoning. In R. Cummins and J. L. Pollock, editors, *Philosophy and AI: Essays at the Interface*, pages 1–22. The MIT Press, Cambridge, Massachusetts, 1991.

[19] M.E. Bratman. *Intention, Plans and Practical Reason*. David Hume Series. CSLI, 1999.

[20] J. Broersen, A. Herzig, and Troquard N. A reading companion to the ESSLLI course "logics for agency and multi-agent systems". Technical report, European Association for Logic, Language and Information, 2007.

[21] J. Broersen, A. Herzig, and N. Troquard. A stit-extension of ATL. In M. Fisher, W. van der Hoek, B. Konev, and A. Lisitsa, editors, *JELIA*, volume 4160 of *Lecture Notes in Computer Science*, pages 69–81. Springer, 2006.

[22] Jan Broersen. A complete stit logic for knowledge and action, and some of its applications. In M. Baldoni, T. C. Son, M. B. van Riemsdijk, and M. Winikoff, editors, *Declarative Agent Languages and Technologies 2008*, volume 5397 of *Lecture Notes in Computer Science*. Springer, 2009.

[23] Jan Broersen, Andreas Herzig, and Nicolas Troquard. From coalition logic to stit. *Electr. Notes Theor. Comput. Sci.*, 157(4):23–35, 2006.

[24] Rodney A. Brooks. *Cambrian intelligence: the early history of the new AI.* MIT Press, Cambridge, MA, USA, 1999.

[25] M. A. Brown. On the logic of ability. *Journal of Philosophical Logic*, 17:1–26, 1988.

[26] M. A. Brown. Action and ability. *Journal of Philosophical Logic*, 19(1):95–114, 1990.

[27] J. Bryson. Where should complexity go? cooperation in complex agents with minimal communication. In W. Truszkowski, C. Rouff, and M. G. Hinchey, editors, *WRAC*, volume 2564 of *Lecture Notes in Computer Science*, pages 303–319. Springer, 2002.

[28] E. Y. Chang, Z. Manna, and A. Pnueli. Characterization of temporal property classes. In W. Kuich, editor, *ICALP*, volume 623 of *Lecture Notes in Computer Science*, pages 474–486. Springer, 1992.

[29] B. F. Chellas. *The Logical Form of Imperatives.* Perry Lane Press, Stanford, California, 1969.

[30] B. F. Chellas. *Modal Logic, an introduction.* CUP, 1980.

[31] B. F. Chellas. Time and modality in the logic of agency. *Studia Logica*, 51:485–517, 1992.

[32] L. Cohen, G. Avrahami, M. Last, and A. Kandel. Efficient learning algorithms for agents mining time-changing data streams. In *CIMCA/IAWTIC*, page 257. IEEE Computer Society, 2006.

[33] R. Conte and C Castelfranchi. *Cognitive and social action.* UCL press, 1995.

[34] Belnap, Jr. N. D. and M. Perloff. Seeing to it that: a canonical form for agentives. *Theoria*, 54:175–199, 1988.

[35] M. Dastani, F. Arbab, and F. de Boer. Coordination and composition in multi-agent systems. In *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 439–446, New York, NY, USA, 2005. ACM.

[36] M. de Rijke. Advances in modal logic startup document, (http://remote.science.uva.nl/ mdr/aiml/aiml-startup.html), 1996.

[37] John Debenham. A multi-agent architecture for process management accommodates unexpected performance. In *ACM symposium on Applied computing*, Como, 2000.

[38] E. W. Dijkstra. On anthropomorphism in science (http://www.cs.utexas.edu/users/ewd/ewd09xx/ewd936.pdf). circulated privately, sep 1985.

[39] R. Erdur and İ. Seylan. The design of a semantic web compatible content language for agent communication. *Expert Systems*, 25(3):268–294, July 2008.

[40] R. Fagin. A quantitative analysis of modal logic. *Journal of Symbolic Logic*, 59(1):209–252, 1994.

[41] R. Fagin and J. Y. Halpern. Reasoning about knowledge and probability. *Journal of the ACM*, 41(2):340–367, 1994.

[42] R Fagin and M. Y. Vardi. An internal semantics for modal logic. In *The seventeenth annual ACM symposium on Theory of computing*, pages 305–315. ACM, 1985.

[43] J. Ferber. *Multi agent systems, an introduction to distributed artificial intelligence.* Addison Wesley, 1999.

[44] J. Ferber and J-P. Müller. Influences and reaction: a model of situated multiagent systems. In *In: Second International Conference on Multi-agent Systems. AAAI*, 1996.

[45] N. Friedman and J. Y. Halpern. Modeling belief in dynamic systems, part II: Revision and update. *Journal of Artificial Intelligence Research*, 10:117–167, 1999.

[46] P. Gärdenfors. *Knowledge in Flux. Modeling the Dynamics of Epistemic States.* NIT Press, July 1988.

[47] M.P. Georgeff and F.F. Ingrand. Decision making in an embedded reasoning system. In *Proceedings of the International Joint Conference on Artificial Intelligence.* IJCAI, 1989.

[48] C. Gershenson and F. Heylighen. When can we call a system self-organizing? In W. Banzhaf, T. Christaller, P. Dittrich, J. T. Kim, and J. Ziegler, editors, *ECAL*, volume 2801 of *Lecture Notes in Computer Science*, pages 606–614. Springer, 2003.

[49] Rod Girle. *Modal Logics and Philosophy.* Acumen, 2000.

[50] N. Halbwachs. *Synchronous programming of reactive systems.* Kluwer Academic Publishers, 1993.

[51] J. Y. Halpern. *Reasoning about Uncertainty.* The MIT Press, October 2003.

[52] J. Y. Halpern, R. Harper, N. Immerman, P. G. Kolaitis, M. Y. Vardi, and V. Vianu. On the unusual effectiveness of logic in computer science. *The Bulletin of Symbolic Logic*, 7(2):213–236, 2001.

[53] J. Y. Halpern and Y. Moses. A guide to the modal logics of knowledge and belief: Preliminary draft. In *Proc. of the 9th International Joint Conference on Articial Intelligence.* Morgan Kaufmann, 1985.

[54] J. Y. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, 1990.

[55] J. Y. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54(3):319–379, 1992.

[56] J. Y. Halpern and R. Pucella. A logic for reasoning about evidence. *J. Artif. Intell. Res. (JAIR)*, 26:1–34, 2006.

[57] S. O. Hansson. Situationist deontic logic. *Journal of Philosophical Logic*, 26(4):423–448, 1997.

[58] S. O. Hansson. Ten philosophical problems in belief revision. *Journal of Logic and Computation*, 13(1):37–49, February 2003.

[59] D. Harel, D. Kozen, and J Tiuryn. *Dynamic Logic*. Foundations of Computing. MIT press, 2000.

[60] D Heckerman. Bayesian learning. In R. A. Wilson and F. C. Keil, editors, *The MIT Encyclopedia of the Cognitive Sciences*, pages 70–72. MIT Press, 1999.

[61] F. Heylighen. The science of self-organization and adaptivity. In L. D. Kiel, editor, *Knowledge Management, Organizational Intelligence and Learning, and Complexity, in: The Encyclopedia of Life Support Systems (EOLSS)*. EOLSS Publishers, Oxford, 2001.

[62] C. A. R. Hoare. *Communicating sequential processes (electronic edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2004.

[63] P. J. 't Hoen and S. M. Bohte. COllective INtelligence with sequences of actions - coordinating actions in multi-agent systems. In N. Lavrac, D. Gamberger, L. Todorovski, and H. Blockeel, editors, *ECML*, volume 2837 of *Lecture Notes in Computer Science*, pages 181–192. Springer, 2003.

[64] P. J. 't Hoen and E. D. de Jong. Evolutionary multi-agent systems. In X. Yao, E. K. Burke, J. L. Lozano, J. Smith, J. J. M. Guervós, J. A. Bullinaria, J. E. Rowe, P. Tiño, A. Kabán, and H-P. Schwefel, editors, *PPSN*, volume 3242 of *Lecture Notes in Computer Science*, pages 872–881. Springer, 2004.

[65] T. Holvoet. Agents and Petri nets. Petri Net Newsletters (49), 1995.

[66] J. F. Horty. *Agency and deontic logic.* OUP, 2001.

[67] J. F. Horty and N. D. Belnap, Jr. The deliberative STIT: A study of action, omission, and obligation. *Journal of Philosophical Logic*, 24(6):583–644, 1995.

[68] N.R. Jennings, K. Sycara, and M. Wooldridge. A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems*, 1(1):7–38, 1998.

[69] S. Kanger. Law and logic. *Theoria*, 38:105–132, 1972.

[70] H. Kargupta, B. Park, D. Hershbereger, and E. Johnson. Collective data mining: A new perspective toward distributed data mining, 1999.

[71] H. Katsuno and A. O. Mendelzon. On the difference between updating a knowledge base and revising it. In P. Gärdenfors, editor, *Belief Revision*, pages 183–203. Cambridge University Press, 1992.

[72] A. Kenny. Human ability and dynamic modalities. In J. Manninen and R. Tuomela, editors, *Essays on Explanation and Understanding*, pages 209–232. D. Reidel Publishing Company, Dordrecht, 1976.

[73] D. N. Kinny, M. P. Georgeff, and J. Hendler. Experiments in optimal sensing for situated agents. In *Proceedings of the Second Pacific Rim International Conferences on Artificial Intelligence.*, pages 1176–1182, 1992.

[74] M. Klusch, S. Lodi, and G. Moro. Issues of agent-based distributed data mining. In *AAMAS* [1], pages 1034–1035.

[75] M. J. Kollingbaum and T. J. Norman. Noa - a normative agent architecture. In G. Gottlob and T. Walsh, editors, *IJCAI*, pages 1465–1466. Morgan Kaufmann, 2003.

[76] U. P. Kulkarni, K. K. Tangod, S. R. Mangalwede, and A. R. Yardi. Exploring the capabilities of mobile agents in distributed data mining. In *IDEAS*, pages 277–280. IEEE Computer Society, 2006.

[77] R. Kurki-Suonio. *A Practical Theory of Reactive Systems: Incremental Modeling of Dynamic Behaviors (Texts in Theoretical Computer Science. An EATCS Series).* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[78] L. Lamport. Proving the correctness of multiprocess programs. *IEEE Trans. Software Eng.*, 3(2):125–143, 1977.

[79] J. L. Leva. Algorithm 712: A normal random number generator. *ACM Transactions on Mathematical Software*, 18(4):454–455, December 1992.

[80] H. J. Levesque and G. Lakemeyer. *The Logic of Knowledge Bases.* MIT Press, 2000.

[81] Z. Li, C. H. Sim, and M.Y.H. Low. A survey of emergent behavior and its impacts in agent-based systems. In *Industrial Informatics, 2006 IEEE International Conference on*, pages 1295–1300, Aug. 2006.

[82] R. Logie, J. G. Hall, and K. G. Waugh. Reactive food gathering. In F. Toni and P. Torroni, editors, *CLIMA VI*, volume 3900 of *Lecture Notes in Computer Science*, pages 406–413. Springer, 2005.

[83] R. Logie, J. G. Hall, and K. G. Waugh. Towards mining for influence in a multi agent environment. In A. P. Abraham, editor, *IADIS European Conf. Data Mining*, pages 97–101. IADIS, 2008.

[84] E. Lorini, N. Troquard, A. Herzig, and C. Castelfranchi. Delegation and mental states. In *AAMAS '07: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, pages 1–3, New York, NY, USA, 2007. ACM.

[85] A. McCallum. Overcoming incomplete perception with util distinction memory. In *ICML*, pages 190–196, 1993.

[86] F. Michel. The IRM4S model: the influence/reaction principle for multiagent based simulation. In *AAMAS '07: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, pages 1–3, New York, NY, USA, 2007. ACM.

[87] R. Milner. *Communication and concurrency.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.

[88] M. Minsky. *The Society of Mind.* Simon and Schuster, 1987.

[89] E. Oliveira. Robots as responsible agents. In *IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, volume 3, pages 2275–2279, 1997.

[90] J. L. Peterson. Petri nets. *ACM Computing Surveys*, 9(3):223–252, 1977.

[91] I. Pörn. *The Logic of Power.* Blackwell, 1970.

[92] I. Pörn. *Action theory and social science : some formal models / Ingmar Porn.* D. Reidel Pub. Co., Dordrecht, Holland ; Boston :, 1977.

[93] G. Priest. *An Introduction to Non-Classical Logic.* Cambridge University press, 2001.

[94] A. N. Prior. *Past, Present and Future.* Oxford University Press, Oxford, 1967.

[95] A. S. Rao and M. P. Georgeff. Modeling rational agents within a BDI-architecture. In A. J. Fikes and E. Sandewall, editors, *proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning, KR91.* Morgan Kaufmann, 1991.

[96] M. Resnick. *Turtles, termites and traffic jams.* MIT Press, 1994.

[97] J Rosenschein and G Zlotkin. *Rules of encounter.* MIT Press, 1994.

[98] C. Rouff, A. Vanderbilt, W. Truskowski, J. Rash, and M. Hinchey. Verification of nasa emergent systems. In *Engineering Complex Computer Systems, 2004. Proceedings. Ninth IEEE International Conference on*, pages 231–238, April 2004.

[99] A. Rubinstein. *Modeling Bounded Rationality.* Zeuthen lecture book series. MIT Press, 1998.

[100] S. Russell. Learning in rational agents. In M. I. Jordan, M. J. Kearns, and S. A. Solla, editors, *Advances in Neural Information Processing Systems*, volume 10. The MIT Press, 1998.

[101] S. J. Russell and P. Norvig. *Artificial Intelligence: A modern approach.* Prentice Hall, 1995.

[102] M. Schut and M. Wooldridge. Intention reconsideration in complex environments. In C. Sierra, M. Gini, and J. S. Rosenschein, editors, *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 209–216, Barcelona, Catalonia, Spain, 2000. ACM Press.

[103] M Schut, M Wooldridge, and S Parsons. On partially observable mdps and bdi models. In *Foundations and Applications of Multi-Agent Systems*, number 2403 in LNAI, pages 243–259. Springer, 2002.

[104] J. R. Searle. *The Construction of Social Reality.* Free Press, January 1997.

[105] Y. Shoham. Implementing the intentional stance. In R. Cummins and J Pollock, editors, *Philosophy and AI*, pages 261–277. MIT Press, 1991.

[106] Y. Shoham and S.B. Cousins. Logics of mental attitudes in ai: a very preliminary survey. In G. Lakemeyer and B. Nebel, editors, *Foundations of Knowledge Representation and Reasoning*, number 810 in LNAI. Springer-Verlag, 1994.

[107] Y. Shoham, R. Powers, and T. Grenager. Multi-agent reinforcement learning: A critical survey. Technical report, Computer Science Department, Stanford University., 2003.

[108] M.P. Singh. Agent communication languages: rethinking the principles. *Computer*, 31(12):40–47, Dec 1998.

[109] S. Strippgen and K. Peters. The other way round! - collaborative communication with agents. In *ACM Agents99*, pages 108–115, 1999.

[110] P. Stütz and R. Onken. Adaptive pilot modeling within cockpit crew assistance. In G. Salvendy, M. J. Smith, and R. J. Koubek, editors, *HCI (1)*, pages 733–736. Elsevier, 1997.

[111] G. Tesauro. Extending q-learning to general adaptive multi-agent systems. In S. Thrun, L. Saul, and Schölkopf, B., editors, *Advances in Neural Information Processing Systems 16.* MIT Press, Cambridge, MA, 2004.

[112] R. H. Thomason. Indeterminist time and truth-value gaps. *Theoria*, 36:246–281, 1970.

[113] R. H. Thomason. Combinations of tense and modality. In D. Gabbay and F. Günthner, editors, *Handbook of Philosophical Logic, Volume II: Extensions of Classical Logic*, pages 135–165. D. Reidel Publishing Co., Dordrecht, 1984.

[114] B. V. Trân, J. Harland, and M. Hamilton. A combined logic of expectation & observation. a generalisation of BDI logics. In J. A. Leite, A Omicini, L. Sterling, and P. Torroni, editors, *DALT*, volume 2990 of *Lecture Notes in Computer Science*, pages 155–172. Springer, 2003.

[115] T. Tulenheimo. Modal logic of time division. In C. Areces and R. Goldblatt, editors, *Advances in Modal Logic*, pages 363–387. College Publications, 2008.

[116] W. van der Hoek, W. Jamroga, and M. Wooldridge. A logic for strategic reasoning. In *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 157–164, New York, NY, USA, 2005. ACM.

[117] H. Van Dyke Parunak and S Brueckner. Entropy and self-organization in multi-agent systems. In *Agents01*, 2001.

[118] H. Von Foerster. *Understanding Understanding: Essays on Cybernetics and Cognition.* Springer, 2002.

[119] D. Weyns and T. Holvoet. A model for situated multi-agent systems with regional synchronization. In *10th International Conference on Concurrent Engineering, Agent and Multi-Agent Systems, CE'03*, pages 177–188. Springer-Verlag, 2003.

[120] D. H. Wolpert, K. R. Wheller, and K. Tumer. General principles of learning-based multi-agent systems. In O. Etzioni, J. P. Müller, and J. M. Bradshaw, editors, *Proceedings of the Third International Conference on Autonomous Agents (Agents'99)*, pages 77–83, Seattle, WA, USA, 1999. ACM Press.

[121] M. Wooldridge. *An introduction to Multiagent Systems*. Wiley, 2002.

[122] M. Wooldridge and N. R. Jennings. Intelligent agents: Theory and practice. *Knowledge Engineering Review*, 10(2):115–152, 1995.

[123] M. Xu. Busy choice sequences, refraining formulas and modalities. *Studia Logica*, 54:267–301, 1995.

[124] M. Xu. Causation in branching time (i): Transitions, events and causes. *Synthese*, 112(2):137–192, 1997.

[125] M. Xu. Axioms for deliberative *stit*. *Journal of Philosophical Logic*, 27(5):505–552, 1998.

[126] A. Yardimci. A survey on use of soft computing methods in medicine. In J. M. de Sá, L. A. Alexandre, W. Duch, and D. P. Mandic, editors, *ICANN (2)*, volume 4669 of *Lecture Notes in Computer Science*, pages 69–79. Springer, 2007.